NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

OPTICAL DETECTION USING FOUR-LAYER SEMICONDUCTOR STRUCTURES

by

David A. Moore

June 2005

Thesis Advisor:	Gamani Karunasiri
Thesis Co-Advisor:	Douglas Fouts

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** June 2005	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE: Optical Detection Using Four-Layer Semiconductor Structures			**5. FUNDING NUMBERS**
6. AUTHOR(S) David A. Moore			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) Space and Naval Warfare Systems Center, San Diego 52560 Hull St San Diego CA 92152-5001			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

The application of a thyristor (a four-layer P1-N1-P2-N2 semiconductor structure) as an optical detector is explored. Based on laboratory experiments which demonstrated that this device produces a pulse-mode output to incident light, the thyristor is investigated by comparing the existing theory of static forward-biased operation to simulation results obtained using ATLAS by Silvaco, Inc. The results include identification of the holding point on the IV curve by simulating the junction potential across each junction as a function of current, and demonstration that impact ionization is not a critical factor in thyristor operation. A series of simulations were performed which show that the thyristor can be optimized for use as a detector by decreasing the emitter efficiencies by decreasing the doping in the P1 and N2 layers, or by increasing the doping in the P2 layer; the switching voltage can be controlled by selecting the doping and thickness of the N1 layer. A detector device was designed to allow further testing of the thyristor detector using the ABN CMOS process from AMI Semiconductor via the MOSIS service. The design of this device is discussed and simulated IV curves are presented.

14. SUBJECT TERMS Thyristor, optical detector, pulse-mode, Silvaco, ATLAS, simulation.			**15. NUMBER OF PAGES** 92
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

OPTICAL DETECTION USING FOUR-LAYER SEMICONDUCTOR STRUCTURES

David A. Moore
Major, United States Marine Corps
B.S.E.E., University of Florida, 1993

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 2005

Author: David A. Moore

Approved by: Gamani Karunasiri
Thesis Advisor

Douglas Fouts
Co-Advisor

John P. Powers
Chairman, Department of Electrical and Computer Engineering

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The application of a thyristor (a four-layer P1-N1-P2-N2 semiconductor structure) as an optical detector is explored. Based on laboratory experiments which demonstrated that this device produces a pulse-mode output to incident light, the thyristor is investigated by comparing the existing theory of static forward-biased operation to simulation results obtained using ATLAS by Silvaco, Inc. The results include identification of the holding point on the IV curve by simulating the junction potential across each junction as a function of current, and demonstration that impact ionization is not a critical factor in thyristor operation. A series of simulations were performed which show that the thyristor can be optimized for use as a detector by decreasing the emitter efficiencies by decreasing the doping in the P1 and N2 layers, or by increasing the doping in the P2 layer; the switching voltage can be controlled by selecting the doping and thickness of the N1 layer. A detector device was designed to allow further testing of the thyristor detector using the ABN CMOS process from AMI Semiconductor via the MOSIS service. The design of this device is discussed and simulated IV curves are presented.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Existing semiconductor photodetectors produce a direct current output in response to incident light. The magnitude of the current output from the device is related to the intensity of the incident light. Various noise mechanisms exist in photodetectors which cause ransom variations in the output current magnitude. The work contained in this thesis explores a novel photodetector mechanism which uses a thyristor (a four-layer P1-N1-P2-N2 semiconductor structure) in a simple RC circuit as an alternative to the photodiode. The demonstrated response of this circuit is a pulse-mode signal which varies in frequency with the intensity of the incident light.

Thyristors have been used for many decades, with the typical application being a switch. Specially designed thyristors, called LASCRs, have been designed to be turned on using photo-generated carriers. The detector application of the thyristor depends, however, on preventing the thyristor from completely turning-on. Much of the existing literature on thyristor theory is intended to better understand how to enhance the turn-on process. A focus of this work is to identify those features which can be used to prevent complete turn-on of the thyristor, thereby enhancing its suitability as a detector.

A simplified thyristor structure was used for simulations performed by ATLAS from Silvaco, Inc, to investigate the behavior of the device at different operating points under forward biasing. Two key operating points, the switching point where the thyristor begins turn-on and the holding point where the turn-on process is completed, are defined using simulation results of the electric field and potential at the *pn* junctions in the device. The switching point is defined as the point where the built-in electric field at the middle junction reaches its maximum strength. This, of course, corresponds with the maximum applied potential at this junction. The holding point is defined as the point where the applied potential is reduced to zero. In other words, the junction potential is equal to the built-in potential of the junction.

Impact ionization is often described as the current multiplication mechanism which gives the thyristor its characteristic shape. In this thesis, simulations and analysis of the electric field in the device show that impact ionization is not a critical factor.

Simulations show that impact ionization does not occur in some thyristors and does not substantially change the IV curves of devices where it does occur.

A series of simulations were performed to explore the effect of variations to the doping concentrations and thicknesses of the thyristor layers. The effort was to identify those relationships which enhance the operation of the thyristor as a detector. It is critical for a thyristor detector not to complete the turn-on process, which is to say that the current must be kept below the holding current. Simulations showed that the holding current for a given thyristor structure can be increased by lowering the emitter efficiency of the two outer *pn* junctions of the thyristor. This can be done by decreasing the doping concentration of the two emitter layers, P1 and N2, or by increasing the doping concentration of the P2 base layer.

Finally, a practical thyristor detector device was designed for fabrication using a CMOS process with a bipolar option. The design and layout for the device are discussed and simulations of the structure are presented. This device, while not an optimal detector structure, will permit controlled experiments to further develop the thyristor detector concept.

I. INTRODUCTION

The purpose of this thesis was to investigate the static operation of thyristors, specifically with the intent to use them as optical detectors. The goal that this thesis supports is a novel silicon photodetector structure based on the thyristor. Although thyristors and other semiconductor photodetectors have existed for many decades, the application pursued herein represents a departure from the standard direct current output of existing semiconductor photodetector devices. The thyristor, when placed in an appropriate circuit, generates a pulsed output that varies in frequency with changes to the incident light intensity.

This chapter gives a brief introduction to the physics of photon absorption, particularly as it is occurs in photovoltaic detectors. This is followed by a synopsis of the experimental results which motivated this work. The second chapter of the thesis explores the structure and physical operation of thyristors. The basic device is described along with an introduction to the existing static theory of operation for the device. The chapter concludes with some results obtained by computer simulations which are relevant to thyristors in general. The third chapter discusses the design and computer simulation of a detector device created specifically for use as an optical detector which will be fabricated using a commercially-available manufacturing process. The final chapter summarizes the results of the thesis and discuss possible areas of future research.

A. PHOTON DETECTION USING SEMICONDUCTORS

A photodetector can be defined as a class of optical detector which detects photons by a measurable interaction between incident photons with the material, resulting in a change in the physical state of the material [1]. The physical property which changes in semiconductors is the free carrier concentration, provided that the photons have sufficient kinetic energy. The energy of a photon E_{photon} is inversely related to its wavelength λ by

$$E_{photon} = \frac{hc}{\lambda} \tag{1.1}$$

where h is Planck's constant and c is the speed of light. When E_{photon} exceeds the energy band gap E_g of the semiconductor material, the photon can be absorbed. Absorption creates an electron-hole pair of free carriers which can be used to conduct electric current.

Semiconductor material can be used to make two varieties of photodetectors. The first is the photoconductor which responds to light of sufficient energy by reducing the resistance to current flow through additional current carriers. The change in resistance is detected by an external biasing circuit. The second, and more common, type is the photovoltaic detector. This uses an internal electric field to sweep the photon-created carriers away before they can recombine. This is the basis for devices such as photodiodes and solar cells which uses the internal electric field found at unbiased or reverse-biased *pn* junctions. An important limitation to photovoltaic detectors is that only those carriers created in a region where an electric field is present are usable. Carriers created away from the electric field will quickly recombine and be of no further use. [1]

As will be seen in the following chapter, a thyristor is nothing more than a *pnpn* structure which consists of three *pn* junctions in series. Under forward bias of the thyristor, one of those junctions is actually reverse biased. This extends the width of the electric field, making that junction more useable as a photovoltaic detector. In this respect, the thyristor is then nothing more than a more complicated structure that performs the same function as a photodiode. However, the nature of the response of the thyristor is entirely different than a photodiode, as will be discussed in the following section.

B. SUMMARY OF PREVIOUS EXPERIMENTAL RESULTS

This section summarizes laboratory experiments by Karunasiri [2] which have demonstrated that a specially prepared thyristor placed in a simple RC circuit produces a pulsed output when illuminated with light. The thyristor used in this experiment was a Motorola MBS4993 which was a relatively small, commercially available device. It was not designed for this application, but rather as a bidirectional switch. To make the device useable for this application, the protective packaging was etched away exposing the actual semiconductor device to incident light. This is not an ideal device on which to conduct further tests. One of the objectives of this thesis is to design a device which will

provide a better testing platform. A general schematic diagram of the experiment circuit is given in Figure 1 [2].

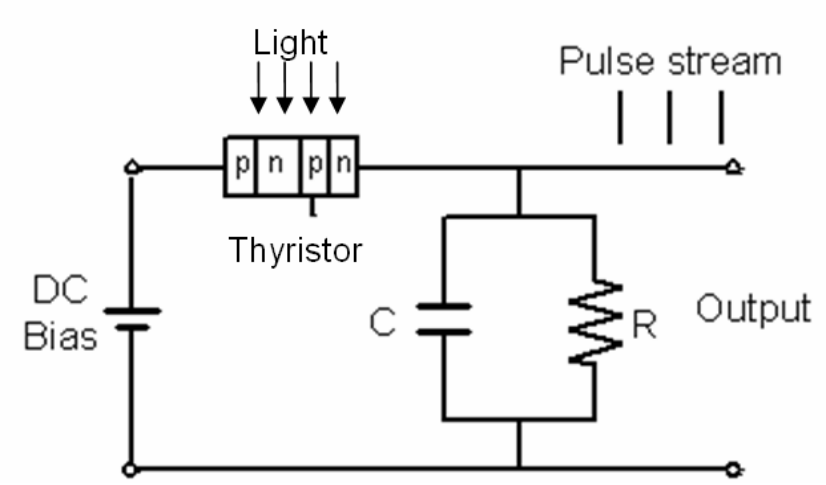

Figure 1. Circuit Diagram (After [2].).

The pulsed output of this circuit is what makes the thyristor detector an interesting concept. The response of a photodiode is a direct current signal which changes amplitude with the intensity of the incident light. The pulse mode output, however, changes frequency with the intensity of incident light.

An example of the pulse mode output is given in Figure 2. Notice that the frequency of pulses increases as the light power on the thyristor is increased.

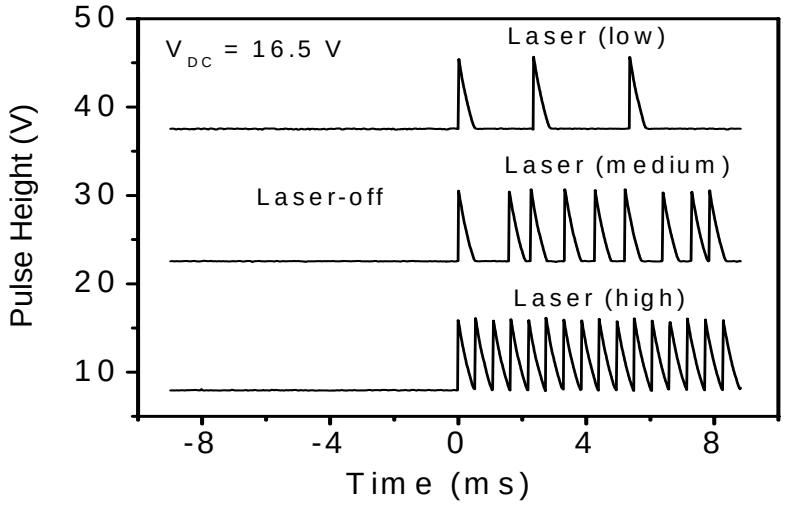

Figure 2. Pulse Mode Output Variation with Incident Light Intensity (From [2].).

There are two aspects of the pulse mode output which make this device worthy of further investigation. The first is related to the potential noise advantage that the pulse mode signal may have over a direct signal. The second is the potential application of this detector. These will be discussed briefly below.

Noise in photodiodes manifests itself in variations in the output current while the incident light is constant. Presumably the same noise sources will be at work in the thyristor photodetector, but the variations will result in changes to the pulse interval. Two of the more significant noise mechanisms in voltaic photodetectors are shot noise and generation-recombination noise. These both involve the statistical variations in the time between events, either between electrons passing across a potential barrier in the case of shot noise or between generation-recombination events in the case of generation-recombination noise, which can be described by Poisson statistics involving extremely large numbers of events. [1]

The Poisson distribution approaches a Gaussian distribution around some mean when the number of events is large, as is usually the case with photons and electrons. The Gaussian distribution is then symmetrical about the mean. This means that the interval between individual events may vary, but the total interval measured over a large number of events will be approximately equal to the number of events times the mean interval. Thus, if the pulse interval is large compared to the interval between electron or photon events, then the statistical variation in event intervals will not significantly impact the pulse interval. [1]

Since noise is not likely to cause significant pulse interval variation shown in Figure 2, another explanation needs to be sought. That work can begin after a better testing platform is fabricated. At this point, it is not thought that the variation represents a problem for use of the thyristor as a detector. This again assumes that the interval distribution is symmetrical about some mean. Then the variation can be mitigated by averaging the interval over a series of pulses.

The second interesting characteristic is the pulse train's resemblance to the waveform used by the human body to transmit electrical signals over nerves. Specifically, the

output of the eye to light stimulation can be described as a “digital spike train” transmitted to the brain [3]. This points to one of the possible future applications of a thyristor-based photodetector, an array of photodetectors to act as a retinal prosthesis.

This chapter has introduced photodetection using semiconductor materials in general, with a specific emphasis on the use of a thyristor to obtain a pulse-mode detector output. The following chapter introduces the structure of the thyristor and static thyristor theory. Simulations were performed to better describe the physical state of the thyristor during forward-biased operation. Finally, the design for a device to be fabricated to permit further testing is introduced.

THIS PAGE INTENTIONALLY LEFT BLANK

II. THYRISTORS

This chapter presents an introduction to the structure and operation of the four-layer semiconductor device known as the thyristor. A brief overview of thyristor structure and static operation is followed by an introduction to static thyristor theory. The chapter concludes with comprehensive static simulations of a thyristor in an effort to better understand thyristor operation.

A. OVERVIEW

1. Structure

The name thyristor is a general term which applies to a large family of four-layer *p-n-p-n* semiconductor devices. A diagram of the basic structure of a thyristor is shown in Figure 3.

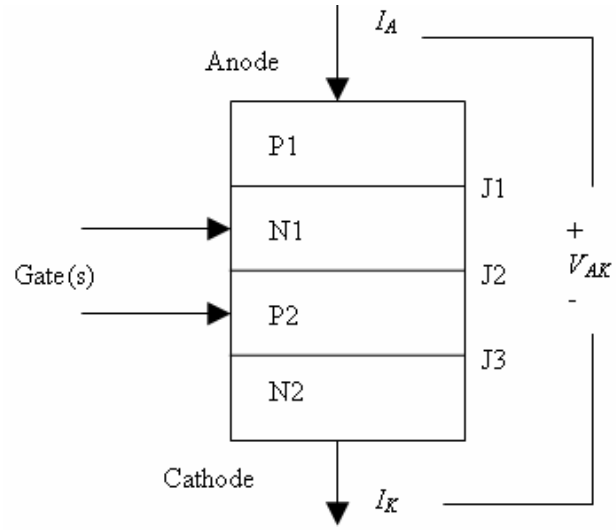

Figure 3. Thyristor Structure.

Currents and voltages described in this thesis will follow the conventions shown in Figure 3. Forward biasing implies that V_{AK} is positive, while positive anode current I_A is defined as positive when flowing into the P1 layer via the anode contact. The literature uses a wide range of names for the different layers and parts of a thyristor. This thesis

will adhere to the names used in Figure 4 and described below. The *anode* is the electrical contact to the P1 layer and the *cathode* is the electrical contact to the N2 layer. *Gate* contacts may be used to either, both, or neither, of the N1 and P2 layers. The junctions formed by the layers are labeled as follows: J1 is the junction between P1 and N1, J2 is the junction between N1 and P2, and J3 is the junction between P2 and N2.

The number of gate contacts is one way of sorting thyristors. For example, a thyristor without gate contacts, that is with only anode and cathode contacts, is known as a Shockley diode; a thyristor with one gate contact to the P2 layer is known as a silicon-controlled rectifier (SCR); and a thyristor with contacts to both the N1 and P2 layers is known as a silicon-controlled switch (SCS). The devices described in this chapter will be Shockley diodes unless clearly specified otherwise.

The relative doping levels of the layers and the widths of the two internal layers are important to thyristor operation. Some general guidelines for these parameters are given by Pierret [4]; the P1 and N2 layers should be highly doped relative to the N1 and P2 layers, and the thicknesses of the N1 and P2 layers should be on the order of a minority carrier diffusion length in each layer. A representative doping profile of a thyristor is shown in Figure 4 where the vertical axis is a logarithmic scale. This plot format is used frequently in this thesis to show various concentration quantities. In each case the vertical scale is the logarithm of the actual concentration value.

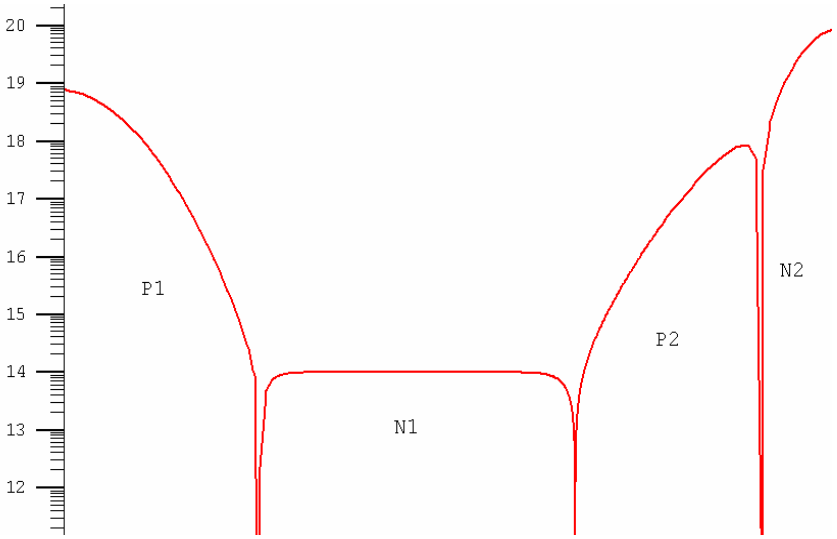

Figure 4. Representative Thyristor Doping Profile (After [4].).

2. I-V Characteristic

Thyristors have a characteristic current vs. voltage, or IV, curve, which is represented in Figure 5. Again, the literature is inconsistent in the labeling of regions and points on the curve. This thesis will consistently use the labels shown in Figure 5.

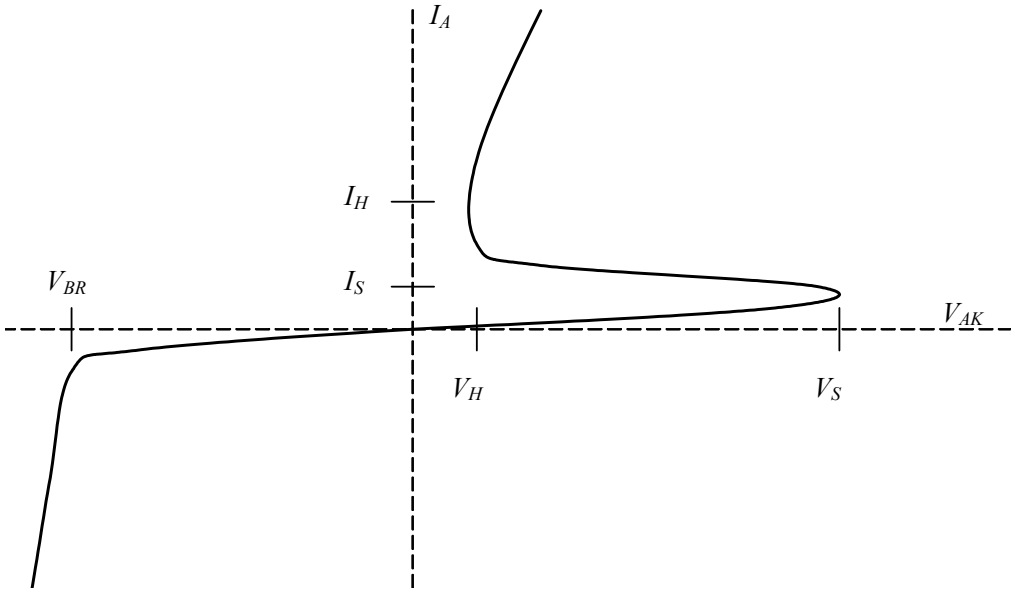

Figure 5. Typical Current-Voltage Curve of a Thyristor (After [5].).

The IV curve can be broken into several operating modes. These operating modes are described by considering the bias voltage V_{AK} and the anode current I_A within the forward and reverse bias regions. Beginning with no bias, as reverse bias $\left(V_{AK} < 0\right)$ is applied, the thyristor initially allows only a small reverse current to flow. This is the reverse blocking mode. Once the bias becomes more negative than the reverse breakdown voltage V_{BR} the thyristor enters the reverse breakdown region. This is characterized by rapidly increasing negative anode current with increasingly negative bias. The reverse bias operation of a thyristor behaves similar to a standard *pn* junction under reverse bias.

The forward bias region is more complicated. Because of the multiple current values, for most voltage points it is easiest to understand this region by considering the anode current rather than the anode-to-cathode voltage. Beginning from zero, increasing the bias voltage initially produces only a very slowly increasing anode current. When the

anode current is less than the switching current I_S, the operating mode is called the forward blocking mode. It is characterized by very high resistance to current. The switching voltage V_S is the maximum voltage across the device. When the current exceeds I_S the device enters the negative resistance mode. As the name indicates, this mode is characterized by negative resistance where increasing anode current results in anode-to-cathode voltage lower than V_S. Once the anode current exceeds the holding current I_H, the device is considered to be *on*. This is the forward conducting mode and it is characterized by extremely low resistance. The holding voltage V_H is the local minimum concurrent with I_H. [5]

It is important to emphasize the distinction between the switching point and the holding point. Switching is not synonymous with turning on. The device begins turning on at the switching point, but is not fully on until the holding point is reached. Thyristors are usually designed to enhance the turning-on process and do not normally operate between the switching and holding points. Returning the device to the forward blocking mode is much more involved from the forward conducting mode than from the negative resistance mode.

3. Switching Mechanisms

There are several ways to induce switching in a thyristor. The most direct is increasing the bias voltage until the anode current exceeds the switching current. This is how a Shockley diode is switched, but it also applies to other thyristors. As long as the biasing source is not current-limited below the holding current, the device will switch on. This mechanism will be used in many simulations throughout this chapter.

Thyristors with a gate contact can be switched using an appropriate gate current which is much smaller than the anode current. This method is typically used in the classic application of the thyristor as a switch. As an example, consider an SCR connected as shown in Figure 6. Recall that an SCR is a thyristor with a gate contact to the P2 layer.

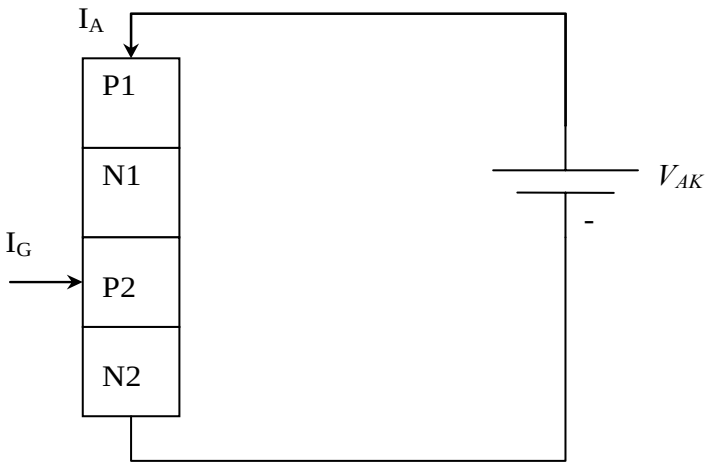

Figure 6. Example SCR Circuit.

Increasing the gate current results in a decrease in the switching voltage V_S as shown in Figure 7 where $I_{G1} < I_{G2} < I_{G3}$. Since the gate current is orders of magnitude smaller than the anode current [4], the thyristor can be used to control very large anode currents and voltages with relatively small gate current. SCR's can be designed to block many thousands of volts in the forward blocking region and to conduct many thousands of amperes in the forward conducting mode [5].

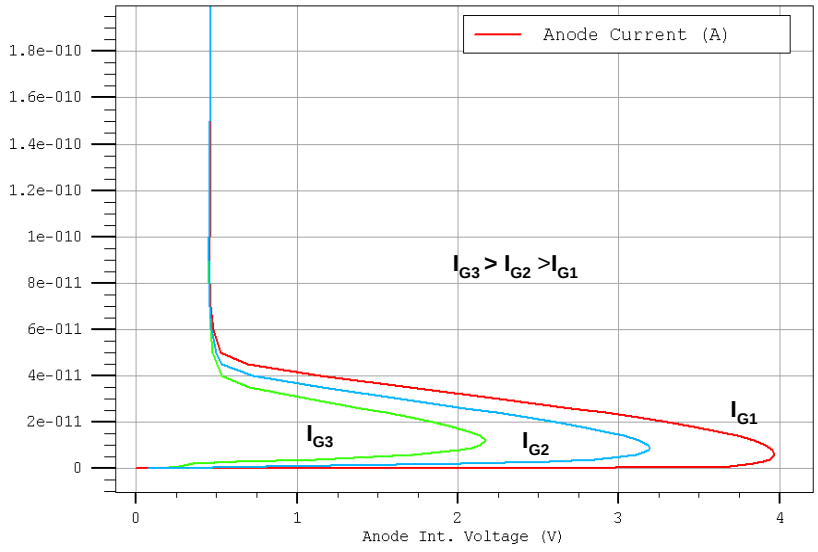

Figure 7. Reduction in V_S with increasing I_G.

Thyristors can also be switched by using incident light. Increasing incident intensity acts much like increasing gate current to lower the switching point, except the carriers are much more quickly swept through the depletion region because they do not have to first travel from the gate. Furthermore, photon absorption generates two carriers, both of which contribute to anode current. [6]

Light activation of thyristor switches allows complete electrical isolation of the switch from the controlling circuit. Devices designed to behave as such are called Light-Activated SCRs (LASCR). LASCRs have a switching point that varies with incident light intensity, as described above. It must be emphasized that LASCRs are not suitable for use as optical detectors since they are designed to turn *on* when sufficient light energy is incident on the device. As mentioned above, turning off a thyristor operating above the holding current level is not a simple process. At best, a LASCR employed as a sensor will be able to indicate that at one point in time light *was* incident on the device. It will not be able to indicate when that may have occurred or even if the light is still incident. As such, it is not of much use as a detector. This points to a design objective for a thyristor used as photodetector; avoid features which enhance the turn-on process.

B. THYRISTOR OPERATIONAL THEORY

This section introduces thyristor operational theory. The focus of this section is limited to the forward-biased region since that is the region of interest in optical detection with thyristors. Coverage of reverse-biased theory can be found in Blicher [7] or Ghandhi [8]. Further, this thesis is interested only in the operation of the thyristor below the holding point. Issues related to completing the turn-on process and turning the device off are also not covered. Again, Blicker and Ghandhi provide good coverage of these topics.

1. Introduction to Thyristor Theory

The development of the thyristor did not lag far behind the development of the bipolar junction transistor. Shockley described what became known as the thyristor in his seminal work on bipolar devices in 1950. He described the four-layer *p-n-p-n* structure as a high current-gain variation of the *p-n-p* transistor, labeling the additional N2 layer as the current multiplication layer [9].

In trying to understand the physical reason for the high gain observed in the high-gain transistor, which occurs in the negative resistance operating region for a thyristor, Shockley effectively ruled out secondary generation, or avalanche multiplication, of carriers. He asserted that the high gain occurs even when electric field strength in the depletion regions is insufficient to induce secondary generation. Shockley concluded his investigation by stating that the likely cause of the high gain is what he calls the *p-n hook theory.* This theory supposes that carrier multiplication in thyristors occurs due to space charge effects of holes. [9]

Figure 8 shows an energy band diagram of a thyristor that is biased slightly below the switching point. The *p-n hook* from Shockley's theory is formed by the potential barrier for holes in the P2 layer. The P2 layer collects holes which were emitted from P1 and carried through N1 by transistor action. In order to understand Shockley's theory, two current components must be considered, electron current coming from N2 and hole current coming from P1. At the switching threshold these currents are beginning to increase because the potential barriers formed by J1 and J3 are lowering as the junctions become forward biased.

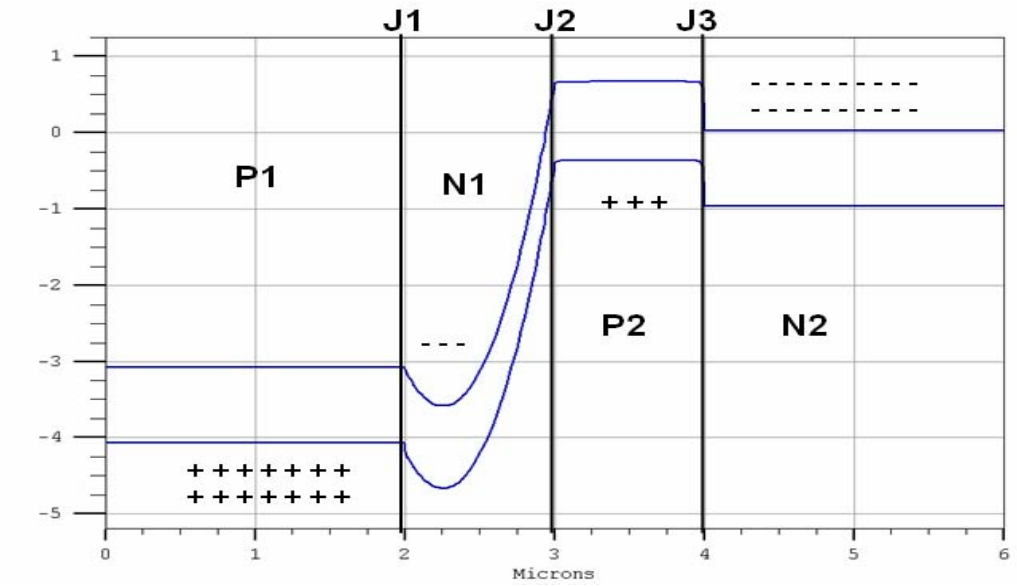

Figure 8. Shockley's *p-n* Hook Collector (After [9].).

The flow of current can be described by following the path of a hole injected from P1 into N1 by thermionic emission. This hole would likely cross the relatively short N1 region without recombining and enter the P2 region by transistor action. The P2 layer represents a low potential area for holes, so they will tend to collect there are excess majority carriers. The holes then impart a net positive charge to the P2 layer which tends to pull in electrons from the N2 layer. Following the reciprocal path, electrons tend to gather in the N1 layer, which tends to pull in more holes. This process is thus regenerative [9].

The relative doping level of N2 is much greater than P2; therefore the current through J3 is dominated by electrons. Since the width of P2 is about one diffusion length, these injected electrons largely transit the P2 layer without recombining. The holes injected through J2 into P2 become excess majority carriers. These holes tend to be trapped in the P2 region because of the potential barrier formed by J3. This results in electron current through J2 that is much larger than the hole current through J3 [9].

Clearly, the current multiplication due to the J3 junction is very important to thyristor operation. However, despite Shockley's early position that this current multiplication is not due to the avalanche process, the idea that avalanche multiplication is involved has persisted. Numerous papers and texts dating from the 1950's-on continue to explain current multiplication using the avalanche process [5, 7, 10, 11]. The issue of avalanche multiplication is revisited later in this thesis.

Other models of thyristor operation have been developed, as well. Perhaps the most useful of these, at least for appreciating the interactions between the layers, is the two-transistor analog. This model is frequently attributed to Moll *et al.* in 1956 [10]. However, the literature shows that in 1952 Ebers mentioned Shockley's claim to be able to get thyristor behavior from two properly connected transistors. Apparently, Shockley established the basis for the model but did not put it in writing [12]. The two-transistor analog is presented in the following section.

2. Two-Transistor Analog

The two-transistor analog considers the thyristor, specifically in this case a Shockley diode, as two Bipolar Junction Transistors (BJTs), one *npn* and one *pnp,* con-

nected as shown in Figure 9. In this figure the two transistors are labeled using both the layers of the thyristor (P1, N1 P2, and N2) and the appropriate BJT layer: emitter (E), base (B) and collector (C). The subscript *pnp* or *npn* is used to distinguish between the two transistors. It must be pointed out that this model is useful only in showing where the device reaches the switching point under forward bias. It does not explain the operation above the switching point. Nevertheless, it is very useful in qualitatively considering thyristor switching.

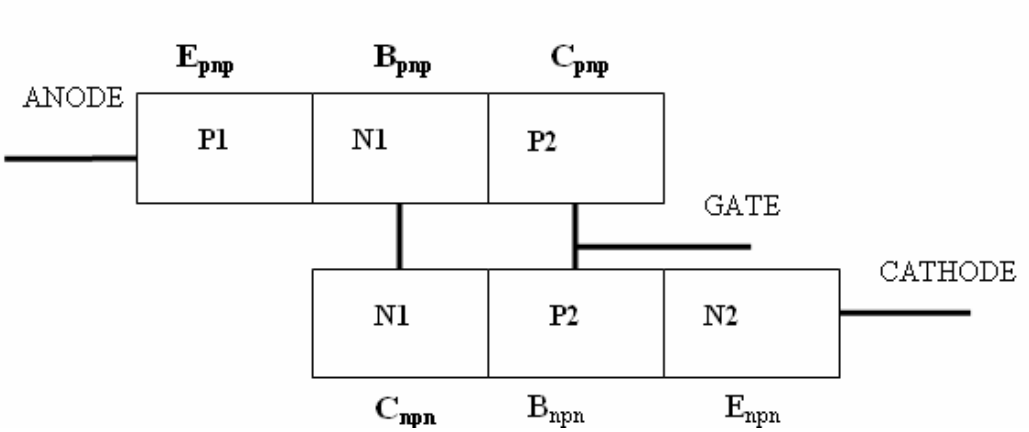

Figure 9. Two-Transistor Analog for the Thyristor.

When forward bias is gradually applied to the device at equilibrium a small current begins to flow. The emitter layers begin injecting holes from P1 into N1 and electrons from N2 into P2 by thermionic emission. Because the N1 and P2 regions are approximately one diffusion width wide, most of these carriers transit their respective BJT bases without recombining and enter in to the collectors. This current tends to forward bias both emitter-base junctions (J1 and J3) and tends to reverse bias the base-collector junctions (J2). In BJT terms, this biases both transistors in the active mode.

The holes gathering in the *pnp* collector (P2) and the electrons in *npn* collector (N1) have no external escape. If these carriers can be seen to move from one transistor's collector to the opposite transistor's base, they become excess majority carriers in those bases. This does not actually involve moving carriers between thyristor layers, only between hypothetical transistors. The excess majority carriers in the base can then diffuse into the emitter regions: P1 for electrons and N2 for holes. This means that holes dif-

fused into N2 will generate electrons injected back into P2, and electrons injected into P1 will generate electrons injected back into N1. This process is then repeated as the minority carriers in the base layers are transmitted into the collectors. Thus the process is regenerative [4]. The switching point in the two transistor model is equated to driving both BJTs into saturation, or forward biasing all *pn* junctions. It will be shown later that this actually happens at the holding point, not the switching point.

The current multiplication that occurs at switching threshold can be understood by considering the relative doping levels between the emitters and the bases. In both cases the emitter is much more heavily doped than the base, or $\mathrm{P1} \gg \mathrm{N1}$ and $\mathrm{N2} \gg \mathrm{P2}$. In a BJT, emitter efficiency γ is defined as the fraction of total current through an emitter-base junction which is due to majority carriers in the base. In the case of J1, this would be given by [4]

$$\gamma = \frac{I_{Ep}}{I_{Ep} + I_{En}} = \frac{I_{Ep}}{I_E} \tag{2.1}$$

where I_{Ep} represents hole current through the junction, I_{En} represents electron current through the junction, and I_E represents the total current through the junction. Another expression for emitter efficiency in terms of material parameters is given by [4]

$$\gamma = \frac{1}{1 + \left(\frac{D_E}{D_B} \frac{L_B}{L_E} \frac{N_B}{N_E} \right) \frac{\sinh(W/L_B)}{\cosh(W/L_B)}} \tag{2.2}$$

where $D_{E,B}$ represents diffusivity in the respective layers, $L_{E,B}$ represents diffusion lengths in the respective layers, $N_{E,B}$ represents doping levels in the respective layers, and W represents the width of the base layer. If (2.2) is reduced to

$$\gamma = \frac{1}{1 + \frac{N_B}{N_E}} \tag{2.3}$$

by assuming the product of the other terms is unity, then γ can be seen to depend on relative doping of the base and emitter. This assumption is made here to demonstrate a point regarding current multiplication. However, in most cases the difference between doping

levels is much larger than the other differences, so doping level difference is usually the primary factor in determining γ[4].

As an example, where P1 is doped at 1×10^{19} cm^{-3} and N1 is doped at 1×10^{19}, (2.3) gives an emitter efficiency of 0.999. Then rewriting (2.1) to solve for I_{Ep} when I_{En} is represented by one electron diffusing from N1 into P1, the result is 999 holes injected from P1 into N1. This example demonstrated the current gain at J1, and a similar process is happening at J3.

Another thyristor explanation has been created which uses the variation in current gain of the two equivalent transistors to define a switching condition. Moll *et al.* define the switching point in terms of the sum of the current gains of the transistors, α_{npn} and α_{pnp}. The thyristor is in the forward blocking mode when $\alpha_{npn}+\alpha_{pnp}<1$ and above the switching point when $\alpha_{npn}+\alpha_{pnp}\geq 1$. The current gain for each BJT is not a constant value, rather it is a strong function of current density. As transistor current increases, as with increasing forward bias, the current gains increase until their sum exceeds unity and the thyristor switches [10]. Moll's explanation is based on a static case. Subsequent work has extended this analysis to the dynamic case [13]. Bear in mind that the values for current gain in the equivalent BJTs are much lower than typical BJT values. The thyristor structure does not correspond with ideal BJT structure, so BJT-based parameters for a thyristor are bound to be less than ideal [4].

3. Junction Capacitance

The model of thyristor behavior presented thus far has not relied upon fundamental semiconductor physics. This is because these models were derived from a qualitative point of view. Nevertheless, some aspects of semiconductor physics can be introduced to better understand the operation of the thyristor. An important property for thyristor operation is junction capacitance.

A *pn* junction has two capacitance components, depletion capacitance and diffusion (or storage) capacitance. Depletion capacitance is normally associated with reverse-biased junctions since it is the only capacitance component under reverse bias. Deple-

tion capacitance also exists in forward-biased junctions, but it is swamped by the other component [14]. The source of depletion capacitance can be explained by considering the charge remaining in a junction at equilibrium as shown in Figure 10. The depletion region consists of two oppositely-charged regions of ionized dopant atoms. Majority free carriers, holes on the left and electrons on the right, are then separated from each other by the depletion width. This separation of opposite charge resembles a parallel plate capacitor.

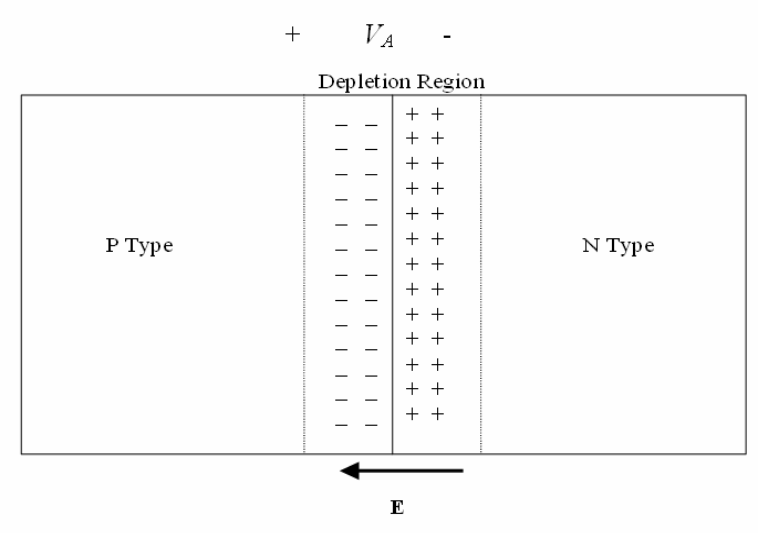

Figure 10. Depletion Capacitance.

As applied voltage V_A across the junction becomes more negative, that is the junction becomes more reverse-biased, the distance between the free charges increases and the capacitance decreases. Junction capacitance for an abrupt junction is given by [14]

$$C_j = A\sqrt{\frac{q\varepsilon N_a^- N_d^+}{2(V_{bi}-V_{AK})(N_a^- + N_d^+)}} \quad \text{for } V_{AK} < V_{bi} \tag{2.4}$$

where A represents cross-section area, q represents elementary charge, ε represents the permittivity of the material, N_a^- and N_d^+ represent ionized acceptor and donor concentrations on either side of the junction, and V_{bi} represents built-in junction potential, which is given by [14]

$$V_{bi} = \frac{k_B T}{q} \ln\left(\frac{N_a^- N_d^+}{n_i^2} \right). \tag{2.5}$$

Equation (2.4) is only valid for applied voltages less than the built-in voltage.

The second capacitance component, known as diffusion or storage capacitance, becomes important as the junction becomes forward biased. As the junction applied bias V_A becomes more positive, the injection of minority carriers across the junction also increases. Diffusion capacitance C_s is due to the rearrangement of minority carrier density [5]. Under forward bias, excess minority carrier concentrations accumulate on both sides of the depletion region with a maximum concentration at the boundary. This creates accumulations of opposite charge separated by a small distance. This is, of course, capacitance. A low-frequency expression for diffusion capacitance at an abrupt junction is given by [5]

$$C_s = A \frac{q}{k_B T} \left(\frac{q L_p p_{no}}{2} + \frac{q L_n n_{po}}{2} \right) \exp\left(\frac{q V_{AK}}{k_B T} \right) \tag{2.6}$$

where $L_{n,p}$ represent diffusion lengths, k_B represents Boltzmann's constant, T represents lattice temperature, and n_{p0} and p_{n0} represent equilibrium minority carrier concentrations.

The relationship between depletion and diffusion capacitances is well defined for reverse-biased junctions, where only depletion capacitance exists, and strongly forward-biased junctions where diffusion capacitance dominates. Less well known is the relationship between the two components for the low forward-biased case. Casey reports that for a junction with built-in voltage of 0.9 V, diffusion capacitance dominates for all applied biases greater than 0.8 V [14]. This is shown graphically in Figure 11. Since diffusion capacitance is negligible unless the junction voltage approaches the built in potential V_{bi}, it can be ignored for low forward-bias and reverse-bias cases. Similarly, junction capacitance increases at a much slower rate than diffusion capacitance, so it can be ignored for strong forward-bias situations.

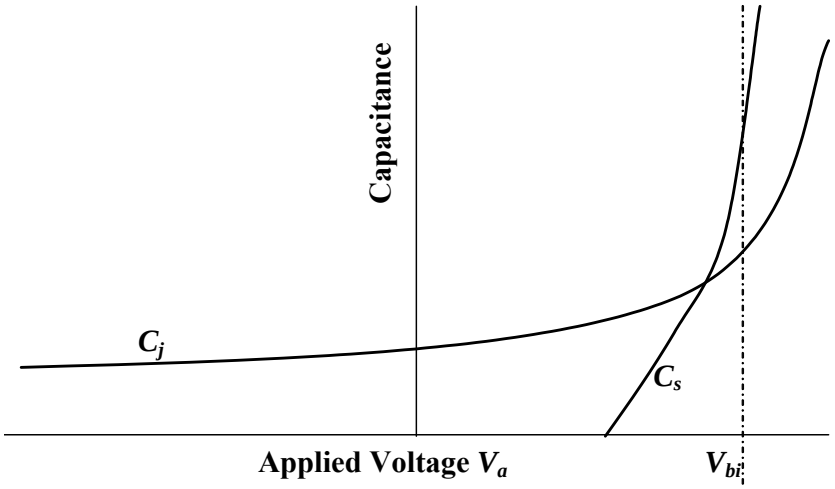

Figure 11. Relation Between Depletion Capacitance and Diffusion Capacitance (After [14].).

This thesis considers only the static case of thyristor operation. The role of junction capacitance in static operation is clearly limited but it must play a great role in dynamic operation. The thyristor itself, since it contains three junctions, can be modeled as three capacitors in series. Moreover, these capacitances change with bias across the device. The interaction of these three junction capacitors, along with the action of the external capacitor shown in the circuit in Figure 1, are likely to be key to understanding the pulsed output from the thyristor detector.

C. SIMULATIONS USING ATLAS BY SILVACO

1. Introduction

ATLAS by Silvaco, Inc. is a powerful, physics-based semiconductor simulator program. It predicts the electrical behavior of charge carriers in semiconductor structures by finding numerical solutions to user-specified carrier transport equations [15]. ATLAS also allows user-selection of a full range of models for physical phenomena such as impact ionization, tunneling, etc. The models used in this thesis are explained in the following section, along with a brief introduction to how ATLAS was used in the rest of this thesis. This is followed by some results obtained from simulations run on a typical thyristor.

2. Physical Models

The first step in performing simulations using ATLAS is determining which models to use. This section describes the models chosen to simulate the thyristor. These models were chosen to simulate relatively small thyristors, with minimum dimensions of about 1 micron, and relatively low switching voltages of about 5V. These models are also expected to be valid for larger thyristors with higher switching voltages but that is beyond the scope of this thesis.

a. Drift-Diffusion Transport Model

The drift-diffusion model of carrier transport explains the motion of charge carriers through drift in an electric field and diffusion from high concentrations to low concentrations. This model is generally sufficient to explain the operation of devices with minimum dimensions greater than one micron [16]. The drift-diffusion model is derived from the more general Boltzmann Transport Equation (BTE).

The BTE describes the time-dependent carrier distribution in six-dimensional position-momentum space. It is given by [16]

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \nabla_r f + \vec{F} \cdot \nabla_p f = C(f) \tag{2.7}$$

where f is the position-momentum carrier distribution function, r and p represent position and momentum, F represents the total force acting on carriers, v represents carrier velocity and C represents the rate of change of the distribution f due to collisions. The BTE can be derived from conservation of carriers in position-momentum space and from carrier trajectories in position-momentum space [17]. Although the BTE gives an accurate representation of carrier distribution, it is not easily solved. In most realistic cases numerous simplifying assumptions must be made. One set of these assumptions leads to the drift-diffusion model.

The drift-diffusion transport model was derived from the BTE using the simplifying assumption that carriers are in equilibrium with the local electric field [17]. The simplest version of the drift-diffusion model is based on two equations which describe the motion of the two carriers due to electric field drift and concentration diffusion, two current continuity equations, and two additional equations describing charge concen-

tration and trap density variation. This results in a set of six coupled partial differential equations.

(1) Carrier Motion Due to Drift and Diffusion. The current densities due to drift and diffusion add together to give the conduction current density *J*. The two equations, with subscript *n* and *p* for electrons and holes, respectively, are given in one dimension as [5]

$$J_n = q\mu_n nE + qD_n \frac{\partial n}{\partial x} = q\mu_n \left(nE + \frac{k_B T}{q} \frac{\partial n}{\partial x} \right) \tag{2.8}$$

$$J_p = q\mu_p pE + qD_p \frac{\partial p}{\partial x} = q\mu_p \left(pE + \frac{k_B T}{q} \frac{\partial p}{\partial x} \right). \tag{2.9}$$

Then the total conduction current is given by $J_c = J_n + J_p$.

(2) Carrier Continuity Equations. The carrier continuity equations are used to account for the flux of free carriers in to and out of an infinitesimal volume. Within this volume, the total rate of carrier concentration increase is the algebraic sum of: the rate of carriers flowing in to the volume, the rate of carriers flowing out of the volume; the rate of carriers being generated within the volume; and the rate of carriers recombining in the volume. The continuity equations are given by [14]

$$\frac{\partial n}{\partial t} = \frac{1}{q} \frac{\partial J_n}{\partial x} + \left(G_n - R_n \right) \tag{2.10}$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q} \frac{\partial J_p}{\partial x} + \left(G_p - R_p \right) \tag{2.11}$$

where $G_{n,p}$ and $R_{n,p}$ represent the generation and recombination rates in the material under consideration. These equations are valid for consideration of both minority carriers and majority carriers. To focus on the minority carriers, which are more important in describing carrier transport in bipolar devices, these equations can be modified to represent minority carrier continuity. To do this, the current density terms in the continuity equation are replaced by the drift-diffusion current density expression. After talking the derivatives (2.10) and (2.11) reduce to [14]

$$\frac{\partial n_p}{\partial t} = n_p \mu_n \frac{\partial E}{\partial x} + \mu_n E \frac{\partial n_p}{\partial x} + D_n \frac{\partial^2 n_p}{\partial x^2} + \left(G_n - R_n \right) \tag{2.12}$$

$$\frac{\partial p_n}{\partial t} = p_n \mu_p \frac{\partial E}{\partial x} + \mu_p E \frac{\partial p_n}{\partial x} + D_p \frac{\partial^2 p_n}{\partial x^2} + \left(G_p - R_p\right). \tag{2.13}$$

In the steady state, defined by $\partial / \partial t = 0$, the electric field term E is independent of time. Therefore the partial derivative of E becomes an ordinary derivative. The change in the electric field strength with respect to distance can be related to the net charge concentration ρ by Gauss's Law [14]

$$\frac{dE}{dx} = \frac{\rho(x)}{\varepsilon} \tag{2.14}$$

where ε represents the dielectric constant. The meaning of (2.14) is that when the electric field in a region is constant, that region must also be balanced in charge. Conversely, any region not in charge balance must add to or detract from the electric field. Thus in a neutral region the first and second terms on the right side from (2.12) and (2.13) are eliminated. The continuity equations in neutral regions simplify to [14]

$$\frac{\partial n_p}{\partial t} = D_n \frac{\partial^2 n_p}{\partial x^2} + \left(G_n - R_n\right) \tag{2.15}$$

$$\frac{\partial p_n}{\partial t} = D_p \frac{\partial^2 p_n}{\partial x^2} + \left(G_p - R_p\right). \tag{2.16}$$

These equations reduce even further in the steady state to

$$D_n \frac{\partial^2 n_p}{\partial x^2} + \left(G_n - R_n\right) = 0 \tag{2.17}$$

$$D_p \frac{\partial^2 p_n}{\partial x^2} + \left(G_p - R_p\right) = 0. \tag{2.18}$$

Finally, in the separate case when carrier distribution is uniform, (2.15) and (2.16) reduce to [14]

$$\frac{\partial p_n}{\partial t} = \left(G_p - R_p\right) \tag{2.19}$$

and

$$\frac{\partial p_n}{\partial t} = \left(G_p - R_p\right). \tag{2.20}$$

(3) Additional Equations. The remaining two equations in the drift-diffusion set are described below. The first expresses conservation of carriers in trap states. It is given by [18]

$$\frac{\partial n_{traps}}{\partial t} = \left(G_p - R_p\right) - \left(G_n - R_n\right). \tag{2.21}$$

In most cases, including the device discussed in this thesis, there is no appreciable time change in the number of trap states, so (2.21) can be ignored [18]. The last expression accounts for the generation of electric field from net charge [19]

$$\nabla \cdot E = \frac{q}{\varepsilon}\left[p - n + N_d^+ - N_a^- \right]. \tag{2.22}$$

This can also be expressed in terms of electrostatic potential V [18]

$$-\nabla^2 V = \rho = \frac{q}{\varepsilon}\left[p - n + N_d^+ - N_a^- \right]. \tag{2.23}$$

This results in a set of six coupled partial differential equations comprised of (2.8), (2.9), (2.10), (2.11), (2.21), and (2.22). This set is usually reduced to five by disregarding (2.21). Semiconductor simulator programs using the drift-diffusion model simply solve these equations numerically at many points in the device. This is the approach used by ATLAS in this thesis [18].

b. Impact Ionization

Impact ionization occurs when free carriers gain sufficient kinetic energy from local electric fields to break covalent bonds of bound electrons upon collision. This process requires a high electric field over a long enough distance to accelerate the carriers to the ionization energy E_i. Furthermore, the distance between collisions must also be large enough to allow acceleration to the ionizing velocity before a collision removes the kinetic energy from the carrier. The general expression for impact ionization rate G is given by

$$G = \alpha_n J_n + \alpha_p J_p \tag{2.24}$$

where $\alpha_{n,p}$ represent ionization coefficients. The ionization coefficients represent the number of electron-hole pairs generated by a carrier per unit length traveled when the kinetic energy of the carrier is greater than E_i [15].

ATLAS allows the user to choose from a number of impact ionization models. These models can be classified into two main types, local models (which are dependent on the local electric field) and non-local models (which depend on the kinetic energy of the carrier).

Local models will normally overestimate the rate of impact ionization [15]. This can be explained by considering the motion of a carrier through an electric field with a sharply-peaked magnitude above the ionizing threshold. This structure is typical of the field at the metallurgical junction of a *p-n* junction. The local model will then calculate the impact ionization across the entire width of the peak. However, since carriers require acceleration in that electric field over some relatively large distance before they actually exceed ionization energy, the local model will overestimate the ionization rate [15]. Non-local models calculate impact ionization by considering carrier kinetic energy rather than electric field strength.

A local-type model was chosen for the initial simulations because of the overestimation tendency of local models. It will be shown later that impact ionization is not strongly involved in small thyristor operation. In certain cases impact ionization does not occur at a high enough rate to change the IV curve of the device. The ATLAS model selected for this simulation calculates the ionization coefficients using

$$\alpha_n = A_N \exp\left(-\frac{B_N}{E}\right) \tag{2.25}$$

$$\alpha_p = A_P \exp\left(-\frac{B_P}{E}\right) \tag{2.26}$$

where $A_{N,P}$ and $B_{N,P}$ are user-definable parameters and E represents electric field strength [15].

c. Mobility

ATLAS also allows numerous mobility models to be used in simulations. Simulations of bipolar devices, such as the thyristor, can typically be accomplished using two mobility models, concentration-dependent mobility and field-dependent mobility. The concentration-dependent model uses a look-up table to assign mobility values based on the total doping concentration at a point. These data have been determined through empirical observations. Several mathematical models have been derived to describe the relationship between concentration-dependent mobility and total doping concentration. One useful mode for silicon at 297 K is given by [14]

$$\mu_n = \frac{1360 - 92}{1 + \left(\frac{N}{1.3 \times 10^{17}} \right)^{0.91}} + 92 \tag{2.27}$$

and

$$\mu_p = \frac{468 - 49.7}{1 + \left(\frac{N}{1.6 \times 10^{17}} \right)^{0.7}} + 49.7 . \tag{2.28}$$

The field-dependent model is used to model carrier velocity, including the velocity saturation effect at higher electric fields. The field-dependent equations are given by [15]

$$\mu_n(E) = \mu_{n0} \left[1 + \left(\frac{\mu_{n0} E}{v_{sat,n}} \right)^2 \right]^{-1/2} \tag{2.29}$$

$$\mu_p(E) = \mu_{p0} \left[1 + \left(\frac{\mu_{p0} E}{v_{sat,p}} \right) \right]^{-1} \tag{2.30}$$

where $v_{sat,\, n,p}$ represents the carrier saturation velocity, and $\mu_{n,p0}$ represents the low-field mobility for the material.

d. Recombination

Two recombination processes were selected for use in these simulations. The first is an impurity concentration-dependent version of the Shockley-Hall-Read (SRH) model. The SRH process generates phonons from electron-hole recombinations which take place using trap levels which exist in the bandgap of real semiconductors. The concentration-dependent version of this model accounts for a decrease in carrier life-times in the presence of moderate or higher doping levels. The carrier lifetimes $\tau_{n,p}$ are determined by [15]

$$\tau_n = \frac{\tau_{n0}}{1 + \frac{N}{N_{SRHN}}} \tag{2.31}$$

$$\tau_p = \frac{\tau_{p0}}{1 + \frac{N}{N_{SRHP}}} \tag{2.32}$$

where N is the total doping concentration. Low-concentration lifetimes $\tau_{n0,p0}$ and parameter $N_{SRHN,P}$ are user-definable, with ATLAS default values of 0.1 μs and 5×10^{16} cm^{-3}, respectively. The low concentration lifetime value is low compared to other authors. Law compiled the research of several researchers and estimated the low-concentration to be 30 μs for electrons and 10 μs for holes [20].

The second recombination process occurs through the *Auger* process. This is a non-radiative process that requires three carriers at the same point in order to occur. Therefore Auger recombination is likely only in highly doped materials [14]. The Auger recombination rate is given by

$$R_A = \mathrm{AUGN}\left(pn^2 - nn_i^2\right) + \mathrm{AUGP}\left(np^2 - pn_i^2\right) \tag{2.33}$$

where n_i represents intrinsic carrier concentration and n and p represent free carrier concentrations. AUGN and AUGP are user definable parameters. The ATLAS default value for n_i is 1.45×10^{10} cm^{-3} in silicon. This value has been disputed by some authors [4, 21]. The results contained herein do not specifically depend on calculations using a precise value for n_i, therefore the ATLAS default value was used throughout.

e. Band Gap Narrowing

Bandgap narrowing is the effect that high doping concentrations can have on semiconductor materials. In silicon this is not observed for concentrations below about 1×10^{18} cm^{-3} [15]. Slotboom and de Graaf derived an expression for bandgap narrowing ΔE_g in silicon BJTs, given by [22]

$$\Delta E_g = 0.009\left(\ln\left(\frac{N}{1\times10^{17}}\right) + \sqrt{\left(\ln\frac{N}{1\times10^{17}}\right)^2 + 0.5}\right). \tag{2.34}$$

The effect of bandgap narrowing is the increase in intrinsic carrier concentration n_i given by

$$n_i = n_{i0}\sqrt{\exp\left(\frac{\Delta E_g}{k_B T}\right)} \tag{2.35}$$

where n_{i0} represents the low doping value for intrinsic carrier concentration.

3. ATLAS Basics

This section discusses the types of simulations used in this thesis and the format of the results of those simulations. In doing so, a thyristor structure is introduced to demonstrate these basics. This structure will also be used in Section D of this chapter which employs ATLAS to further investigate thyristor operation.

a. Input Deck

The basis for any simulation in ATLAS is a file called the *input deck*. The input deck consists of the following five sections: structure specification, material models specification, numerical method specification, solution specification, and results analysis. The details of writing the input deck are beyond the scope of this thesis. The reader is referred to the ATLAS User's Manual for detailed description on input deck syntax [15]. Appendix A contains the input deck used to create the device in Chapter II. Figure 12 shows the hypothetical thyristor used through the rest of this chapter.

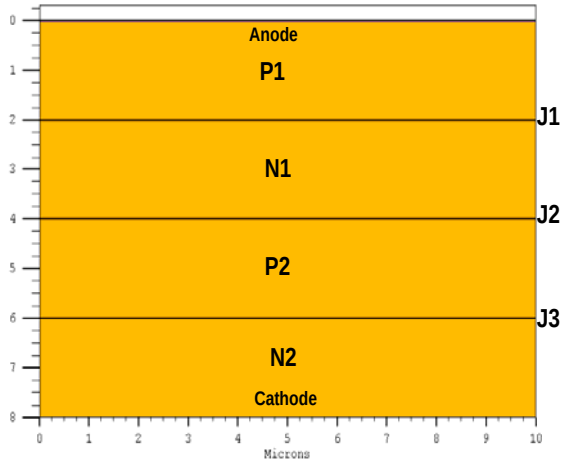

Figure 12. Hypothetical Device and ATLAS Input Deck.

The thyristor used in the initial simulations was a simplified, non-realistic structure. The doping levels and layer thicknesses used in this notional device were established by combining example levels provided in the literature [4] which were refined through multiple simulation runs to produce a uniformly doped layered device with a relatively low switching voltage V_s. The non-physical aspect here is the strictly uniform

doping profile shown in Figure 13. This ideal step-function is not achievable in actual devices, but allows for more clear illustration of the physical conditions here.

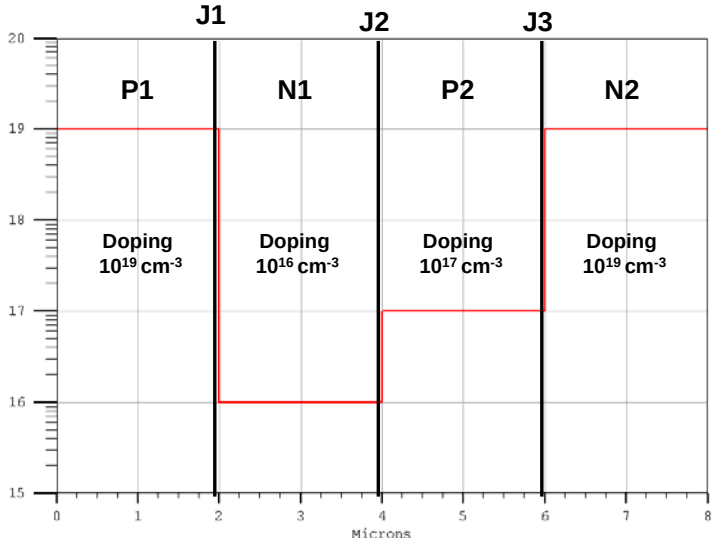

Figure 13. Doping Profile and Layer Thicknesses.

All the simulations performed in this thesis were done in two dimensions. This means that many numerical results, particularly those involving currents, must be considered as quantity per micron of length in the third dimension. This dimension would be into the page. For example, if the current output from a simulation is given as 100 pA, this really means 100 pA per μm of length into the paper. A 100-μm long device with the structural cross-section given produces 100 times the current that the two-dimensional model predicts.

b. ATLAS Outputs

There are two standard output files which come from a completed simulation in ATLAS, the log file and the structure file. The log file is used to display one quantity against another. The best example of this is the current-vs.-voltage curve, or IV curve. Some of the other quantities which can be displayed in a log file display include carrier concentrations, carrier mobility, carrier temperature, carrier generation rate and, of course, voltage and current. The key to understanding the log file display is that the output gives the specified quantity at a specified point as another quantity is varied. This point can be located anywhere inside or on the surface of the device.

In the case of the IV curve, the anode current is measured as the anode bias varies. Another example would be measuring carrier generation in the depletion region of a photodiode as light intensity varies. The simulated forward-biased IV curve for the thyristor described in Figure 13 is shown in Figure 14. The key IV parameters are switching voltage V_S of 16.7 V, switching current I_S of 1 pA, holding voltage V_H of 0.41 V, and holding current I_H of 14 pA. These points are found using the conventional definition of the point where the tangent is vertical, or $dV/dI = 0$. Again, the current figures are per unit micron in the unspecified third dimension.

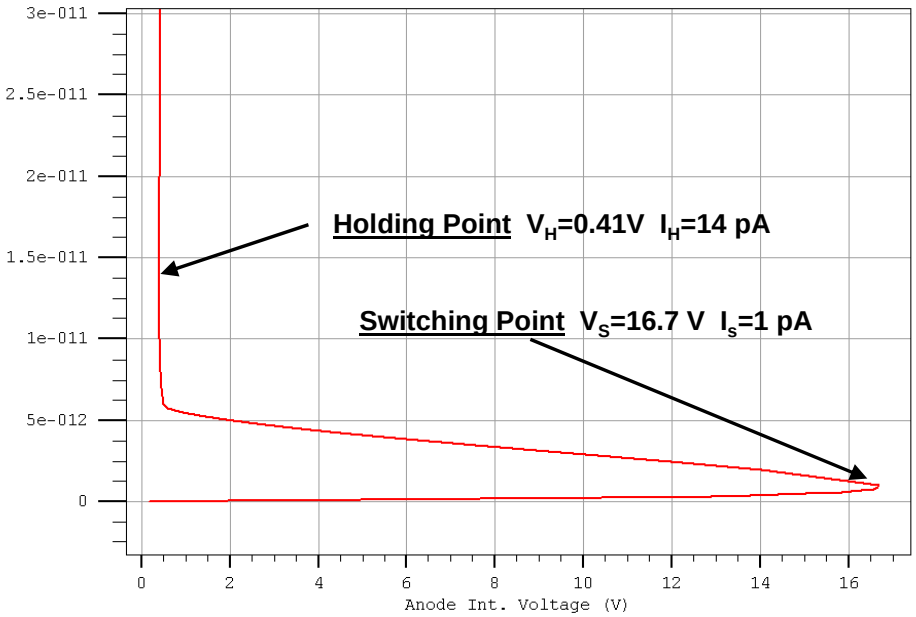

Figure 14. Example Log File Display of Thyristor IV Curve.

The structure file displays a specified quantity in all regions of a device at a specified operating point. A good example of this would be showing the electron distribution at thermal equilibrium. Other quantities that can be shown in a structure file display include electric field, current density, band structure, and many others. The primary structure file display is a multi-color plot where variations are shown by changes in colors. The structure file display also includes a feature called the *cutline*. This feature allows a more accurate numerical-scale display of the desired quantity along a specified line through the device. An additional feature of the cutline display is that multiple quan-

tities can be displayed on the same plot. These features are demonstrated in Figure 15 which shows a basic structure file display along with the cutline display. The cutline, which is taken along a vertical line through the device and plotted on a horizontal line, shows the donor concentration and the electron concentration on a logarithmic scale. The plot shows that the logarithm of the electron concentration in p-type regions approximately agrees with that predicted by the *pn* product [14]

$$n = \frac{n_i^2}{p}. \tag{2.36}$$

Figure 15 also shows that electrons move away from high concentrations at the boundary with a low concentration, as is expected via diffusion.

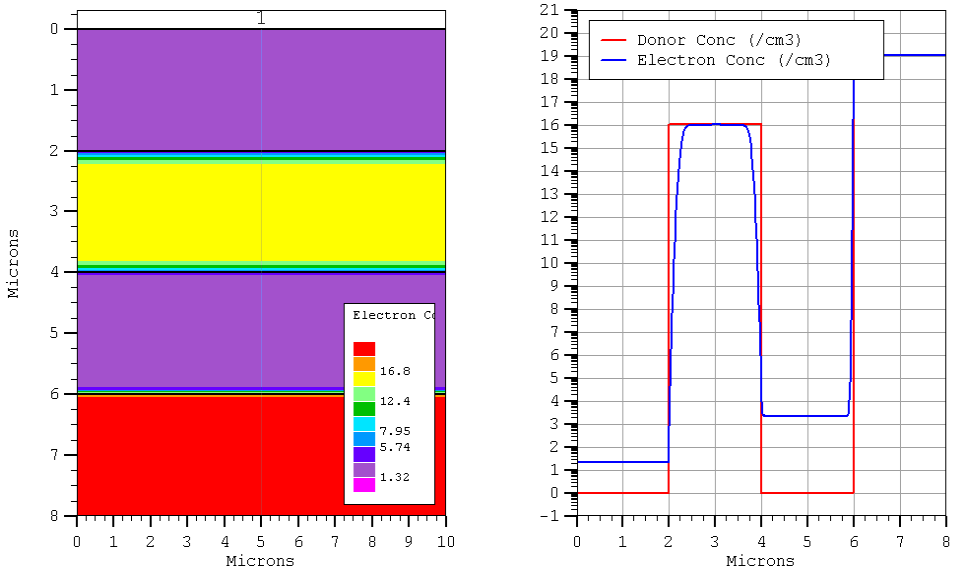

Figure 15. ATLAS Structure Display and Corresponding Cutline Display.

c. Simulation Overview

ATLAS allows the user to choose from a range of simulation types, such as DC, AC, or transient, and to perform them under a diverse range of operating conditions such as bias, current, temperature, and incident light. The simulations contained in this thesis are generally DC analyses conducted both with incident light and without. In each case the temperature is assumed to be 300 K.

Light in ATLAS can be modeled either as a single wavelength source or a multi-spectral source. The model used in this thesis has the same shape as the Air Mass Zero (AM0) solar spectrum [23]. The intensity of the spectral components varies with the overall intensity, which is specified in W/cm^2. This light model was used for convenience. Future work should include characterization using a intensity distribution which is uniform over the desired wavelength range. The relative intensities of the spectral components used for simulations in this thesis are shown plotted in Figure 16.

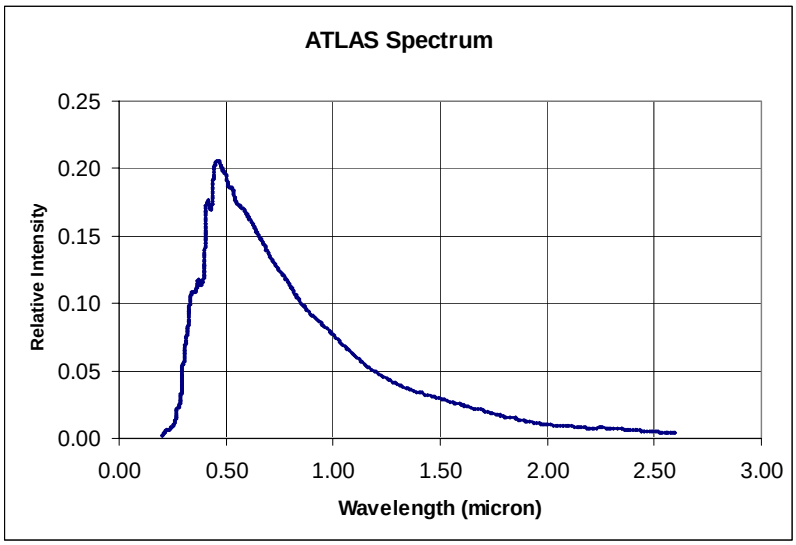

Figure 16. Relative Intensity of Spectral Components.

This section has introduced the ATLAS simulation program and how it will be used in the remainder of this thesis. The models of physical processes assumed to be at work in the thyristor were presented along with how those models are presented in ATLAS. The following section uses ATLAS to explore thyristor operation in detail and to compare the results with operational theory.

D. SIMULATION RESULTS

This section analyzes the results of ATLAS simulations conducted on the device shown in Figure 12. These results, along with analysis, are used to observe and describe changes in the physical state of the device at different operating points.

1. Switching Conditions

This section briefly shows the results of the simulations on the example thyristor in the vicinity of the switching point. The goal here was to establish the physical conditions which occur at this point. Figure 17 shows the change in carrier concentration when biased at the switching threshold.

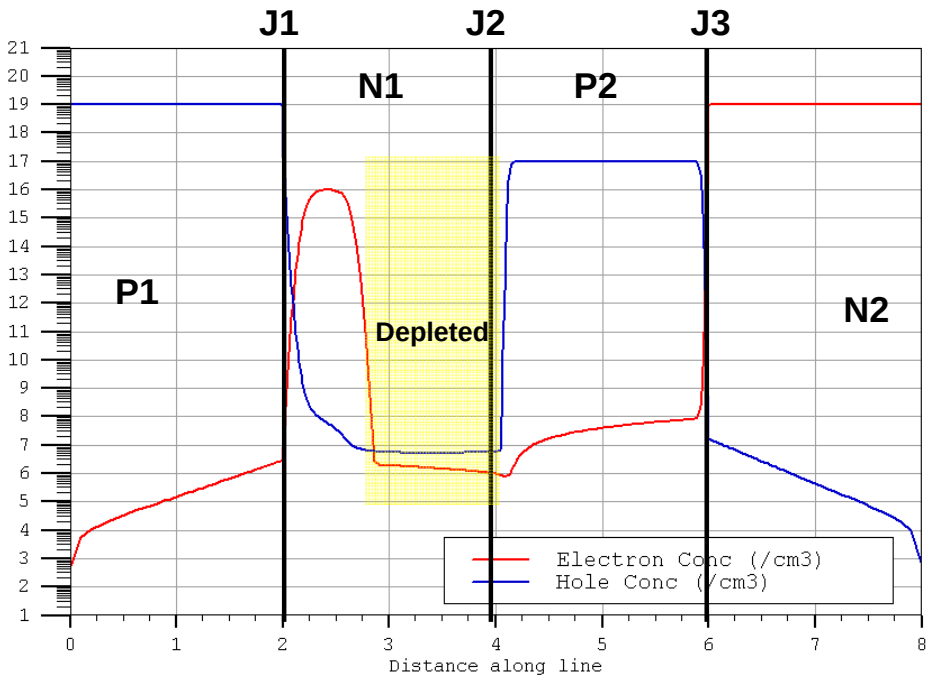

Figure 17. Carrier Concentrations (cm^{-3}) at Switching Threshold.

Qualitatively, Figure 17 describes exactly what is expected at the junctions. J1 and J3 are forward biased, as can be seen in the accumulation of minority carriers at the forward-biased junction. For example, the increase in electrons in P1 and the increase in holes in N2, both near J1, are characteristics of a forward-biased junction. Conversely, the decrease in majority carriers near J2 shows that this junction is reverse biased. This depletion of carriers occurs as the electric field developed at J2 spreads further into N1

and P2. Notice that the carriers are depleted well into the N1 region at switching threshold.

The switching threshold represents the maximum bias across the thyristor, and J1 and J3 are becoming more forward biased as current increases. This means that the junction potential at J2 should also be a maximum at the switching point. This is shown in Figure 18 which plots the current against the voltage across J2. Note that since the voltage is "double valued," the data was generated in ATLAS by increasing the current over the desired range and computing the resulting voltage. This plot format of current vs. voltage is used throughout this thesis.

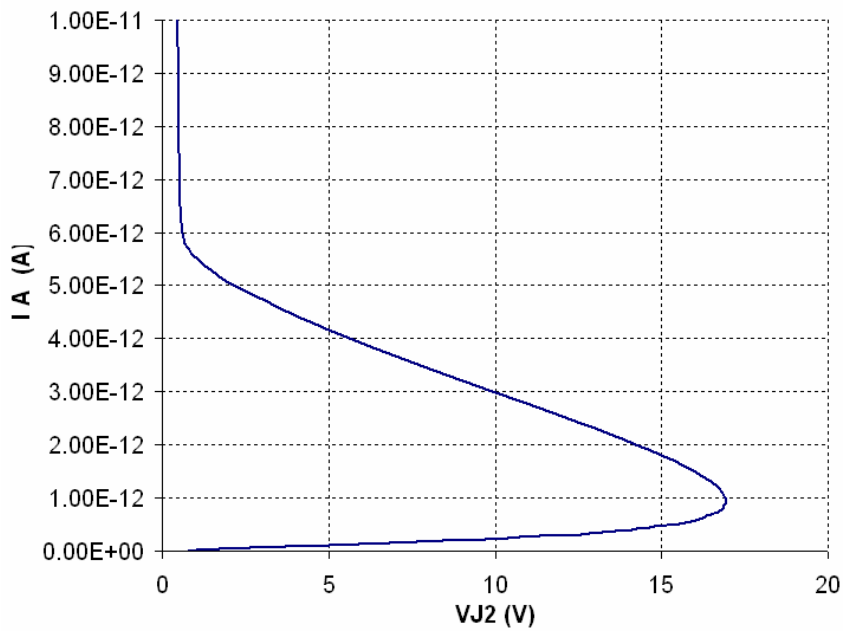

Figure 18. Junction Voltage Across J2 as Current Increases.

The maximum reverse bias across J2 also means that the electric field across J2 will be a maximum. The electric field between J1 and J2 is shown in Figure 19. This shows that the electric field reaches a peak strength of 2.1×10^5 V/cm at J2. The electric field extends farther into N1 because of the lower doping level than in P2.

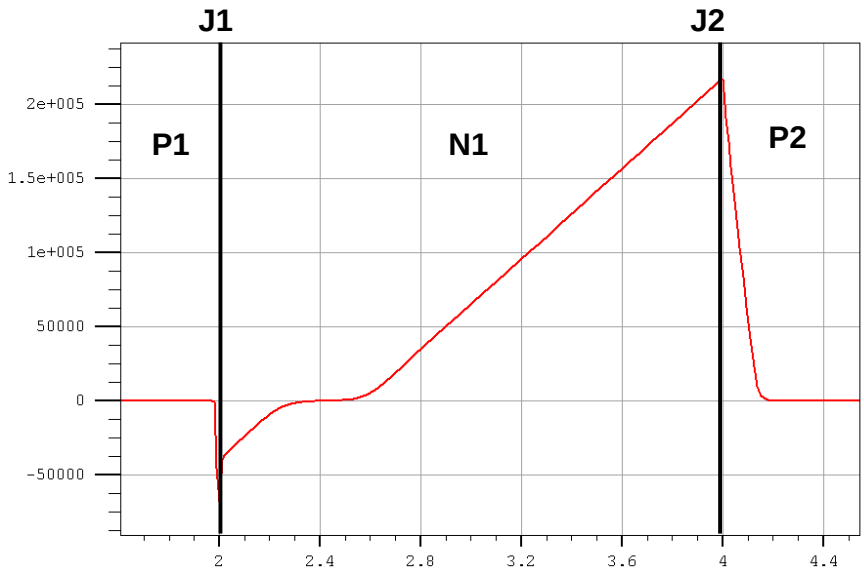

Figure 19. Electric Field in the N1 Layer.

Figure 20 shows the voltages across the junctions J1 and J3. These plots confirm that the junctions are becoming more forward biased as current through the thyristor increases.

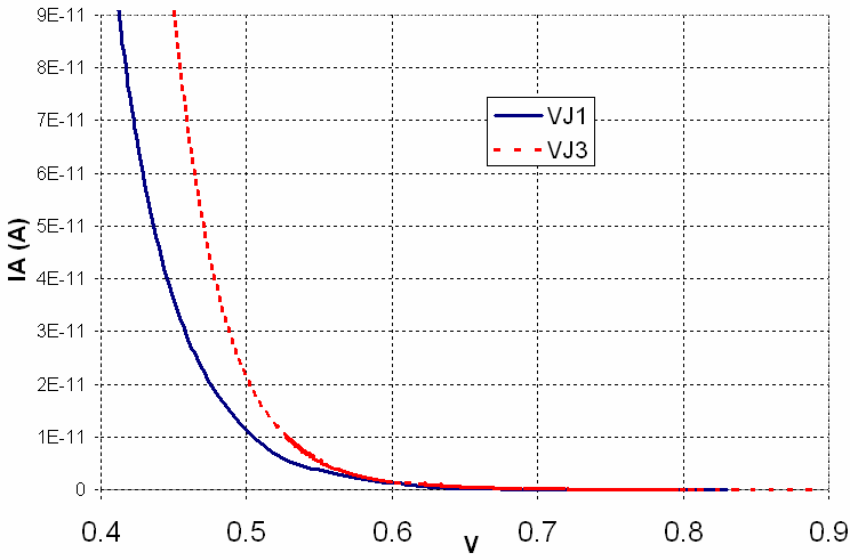

Figure 20. Junction Voltage Across J1 and J3 as Current Increases.

2. Impact Ionization

This section revisits the role of impact ionization in thyristor operation. Recall that Shockley asserted very early on that the current multiplication in a thyristor was not

due to the avalanche process. His reasoning was that the current multiplication was observed at low enough bias voltages to make the avalanche rate extremely low. Despite Shockley's early position that this current multiplication is not due to the impact process, the idea that avalanche multiplication is involved has not diminished. The simulation results in this section will show that while impact ionization *may* occur in some thyristors, it does not occur in all devices.

First, ATLAS simulations are used to show that thyristors can operate without impact ionization. The simulated structure used here uses the same doping profile as used previously, but the thicknesses of N1 and P2 are set to 1 μm. The doping profile is shown in Figure 21.

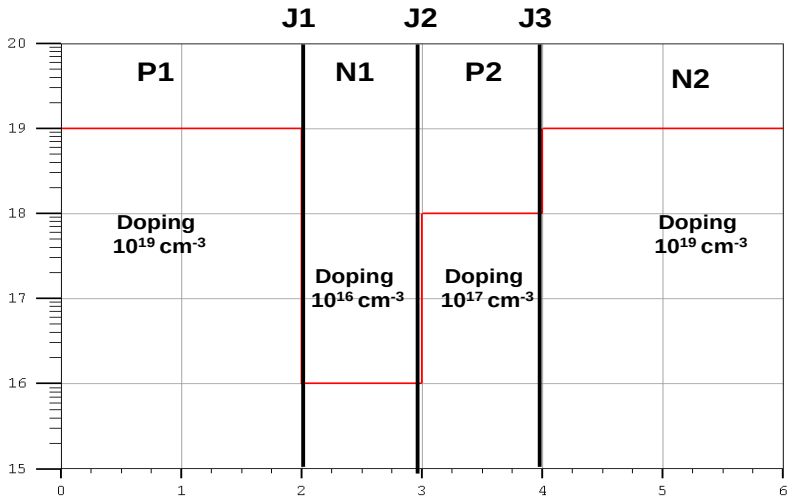

Figure 21. Impact Ionization Demonstration Doping Structure.

ATLAS was then used to generate the simulated IV curve for this device, once using the default impact ionization parameters and once setting those parameters to zero. The later case is equivalent to turning of the impact ionization model. Figure 22 shows the two IV curves. The curve on the right was generated using the impact ionization model with the default parameters used for (2.25) and (2.26), which are described below. The curve on the left was generated with these parameters set to zero. The curves are identical.

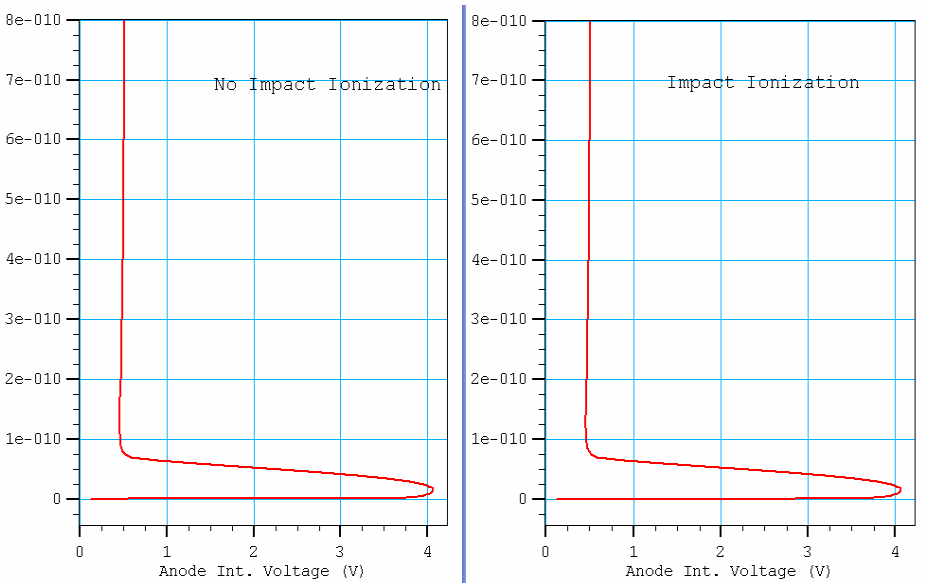

Figure 22. IV Curve Comparison, With and Without Impact Ionization Modeling.

In order to investigate this a bit further, the maximum electric field strength for this device was found to be 1.1×10^{5} V/cm at J2, as shown in Figure 23.

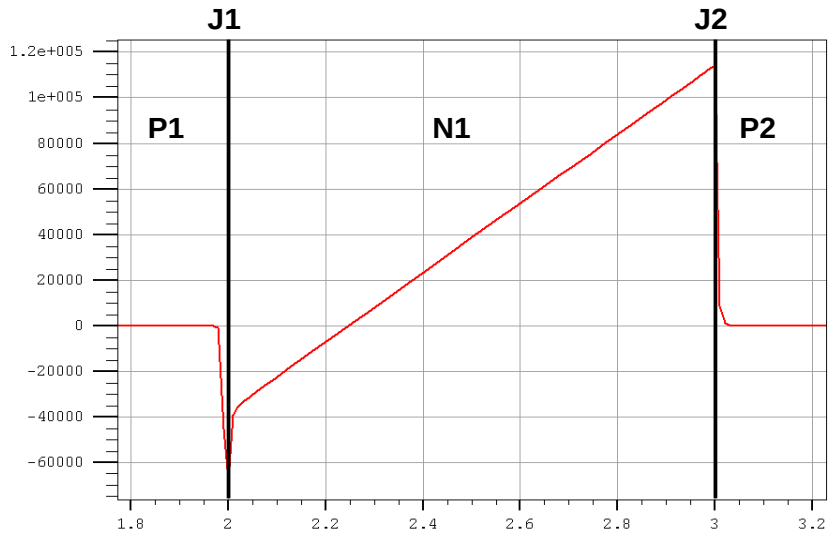

Figure 23. Electric Field Through the N1 Layer.

The mathematical models for impact ionization used in these simulations, which are repeated here from (2.25) and (2.26), are given by

$$\alpha_n = A_N \exp\left(-\frac{B_N}{E}\right) \tag{2.37}$$

$$\alpha_p = A_P \exp\left(-\frac{B_P}{E}\right) \tag{2.38}$$

where E represents the local electric field magnitude. The default parameter values are $A_N = 7.03\times10^5$ cm^{-1}, $B_N = 1.23\times10^6$ V/cm, $A_P = 7.03\times10^5$ cm^{-1}, and $B_P = 1.69\times10^6$ V/cm. The parameter $A_{N,P}$ can be interpreted as the maximum number of carriers that can be generated per unit distance in a *very strong* electric field (i.e., $E \gg B_{N,P}$) while $B_{N,P}$ represents a scaling factor related to where the electric field becomes very strong [24]. Equations (2.37) and (2.38) can then be interpreted as meaning that, unless the electric field E approaches the values given by B_N or B_P, the exponential term will dominate and drive the ionization rate very low. Comparing the maximum electric field in this device to these parameters indicates that this model will show a very low rate of impact ionization.

The best values to use for the parameters $A_{N,P}$ and $B_{N,P}$ are uncertain. However, Maes *et al.* [24] conducted a review of the literature related to impact ionization in silicon. Their results, while inconclusive as far as actual parameter value valid for low field strengths, indicate that the rate of impact ionization would be very low in this device. Using the simulated peak field strength of 1.1×10^5 V/cm, Figures 24 and 25 indicate the values for the ionization rates α_N and α_P would be 2 cm^{-1} and less than 1 cm^{-1}, respectively. Again, this indicates that impact ionization is not a factor in the operation of this thryristor.

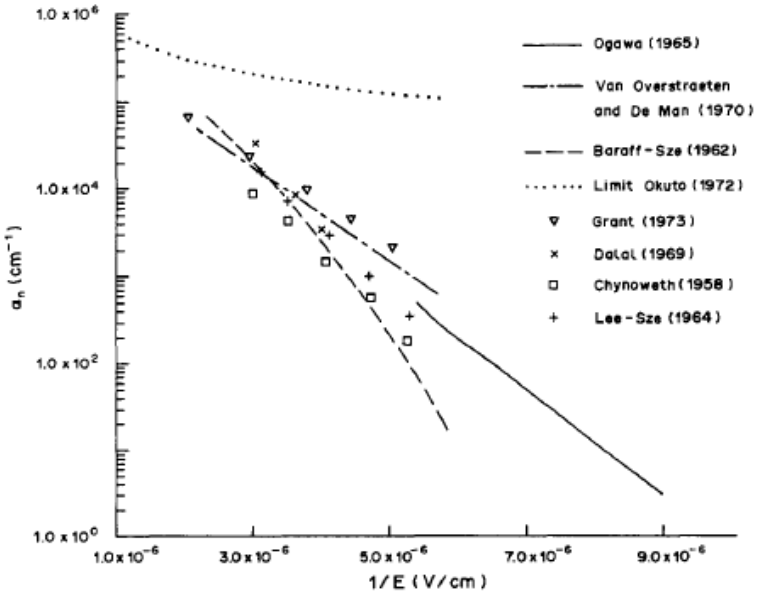

Figure 24. Log of Impact Ionization Rate for Electrons (from [24]).

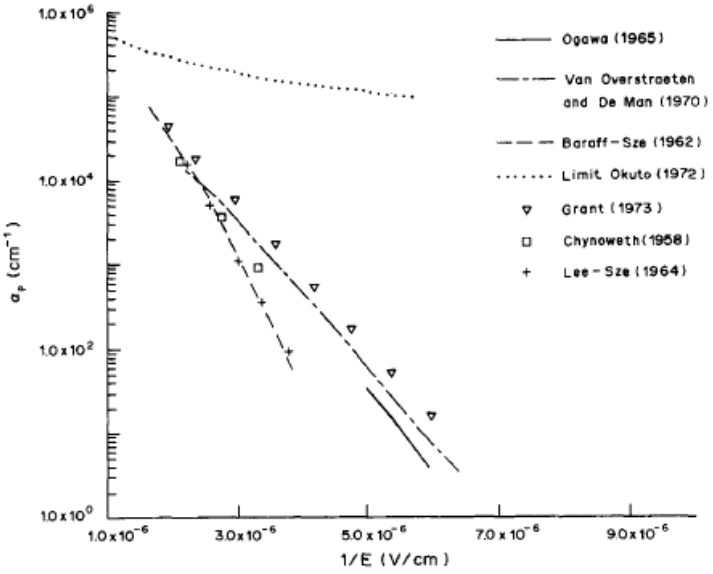

Figure 25. Log of Impact Ionization Rate for Holes (from [24]).

The previous example shows that a thyristor can be simulated to operate properly without impact ionization. However, this does not mean that impact ionization does not occur, just that it is not the primary current multiplication mechanism at work. As an example of a device where impact ionization does have some impact on the IV curve, consider the device introduced in Figure 12 at the beginning of this chapter. The maximum

electric field strength from Figure 19 is 2.1×10^5 V/cm, which is still well below the strong field threshold set by parameters A_N and A_P of 7.03×10^5 cm^{-1}. As Figure 26 shows, the IV curve does show a small effect between simulations with the impact model enabled and disabled.

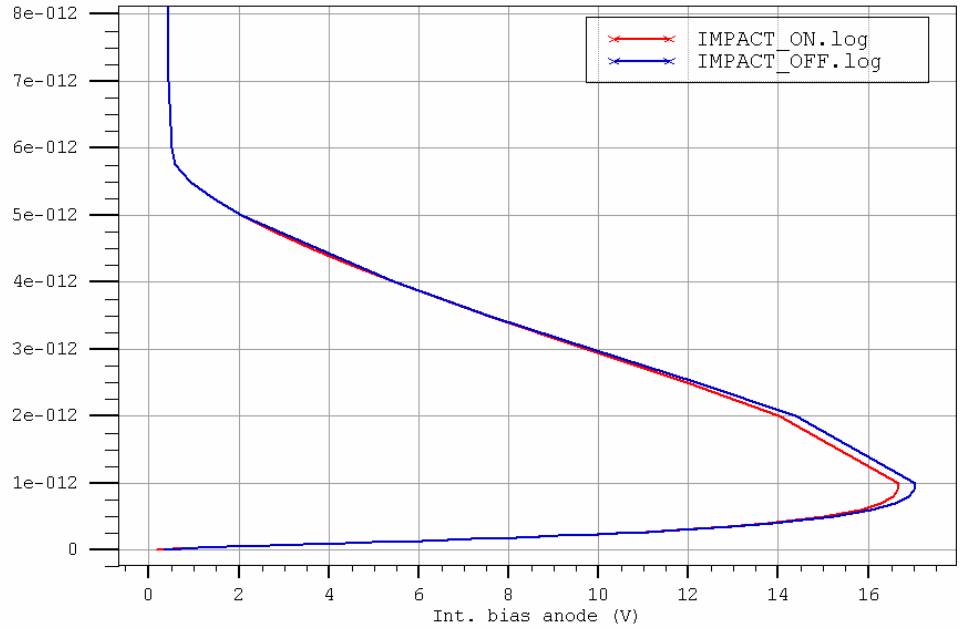

Figure 26. IV Curve Showing Impact Ionization Effects.

The result of this section is that impact ionization is not the primary cause of the current multiplication in a thyristor. While impact ionization may occur in a thyristor, it is not critical to the operation of the device.

3. Holding Conditions

Recall that the two-transistor model for thyristor operation described the switching point by finding where both BJT's become saturated. It was previously shown that this does not happen at switching but can be used as criteria for establishing the holding point. At some anode current value, J2 must become less reverse biased than it is at equilibrium. This can be carried out by ATLAS using two approaches.

The first approach is a graphical analysis of structure plots of the electric field near J2. As J2 becomes slightly forward biased, the electric field will be of a slightly lower magnitude and of slightly smaller lateral extent into the N1 layer. Figure 27 shows

the electric field near J2 simulated at equilibrium and at the holding point as defined in Figure 14. This holding point definition was determined by finding the point on the IV curve with a vertical tangent; in Figure 14 it was found to be 14 pA.

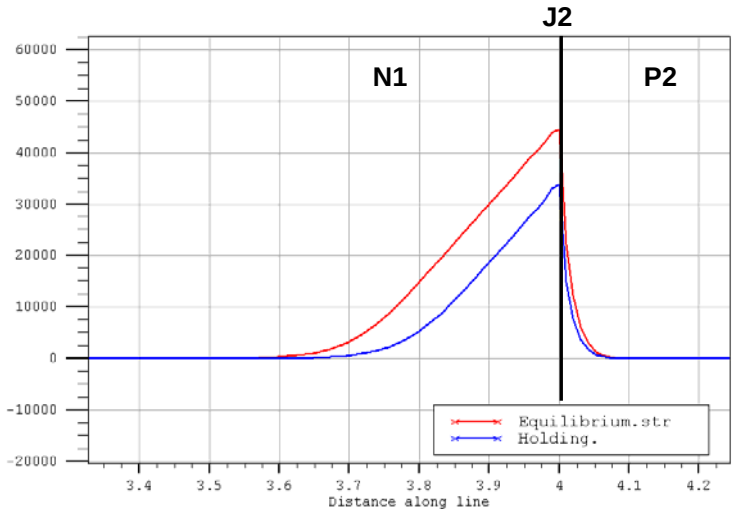

Figure 27. Electric Field Near J2 at the Holding Point.

Figure 27 clearly shows that the electric field at the holding point, shown in blue, is significantly less than at equilibrium, shown in red. This indicates that J2 becomes forward biased at a lower anode current than that defined by finding the vertical tangent to the IV curve.

Further simulations were performed until a more accurate value could be determined to define when J2 becomes forward biased. Figure 28 shows the electric field near J2 again, but this time the blue line represents an anode current of 5.7 pA. In this case, the electric field has a slightly lower peak magnitude and extends slightly less distance into N2 than at equilibrium. This is a better definition for the holding point than the vertical tangent to the IV curve.

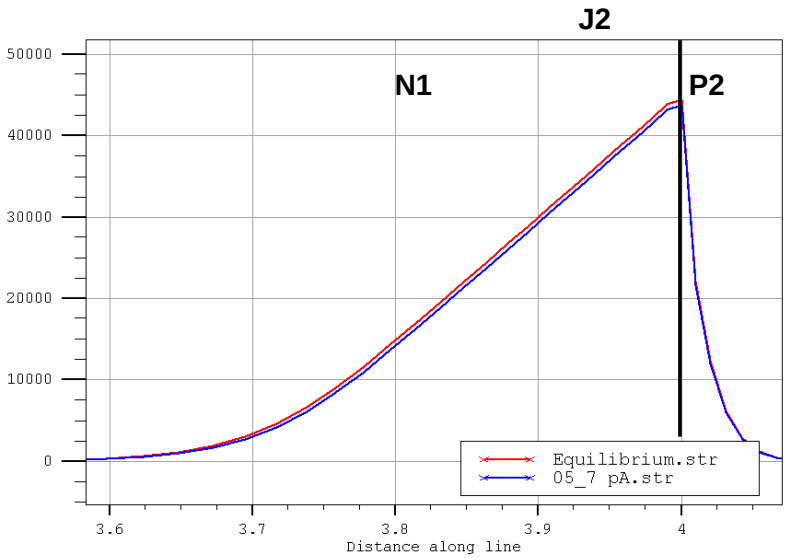

Figure 28. Electric Field Near J2 at the Holding Point.

The second method of finding where J2 becomes forward biased is to find the anode current where the junction potential across J2 (V_{J2}) is equal to the built-in potential V_{bi} across the junction. The built-in potential of a junction was given by (2.5), which is repeated here for convenience

$$V_{bi} = \frac{k_B T}{q} \ln\left(\frac{N_a^- N_d^+}{n_i^2}\right). \tag{2.39}$$

Equation (2.39) applied to J2 gives a built-in voltage of 0.756 V. Figure 29 shows a different view of the data contained in Figure 18 with this built-in potential plotted to determine the corresponding anode current. This value, indicated by the dot, is approximately the same as determined using the prior method.

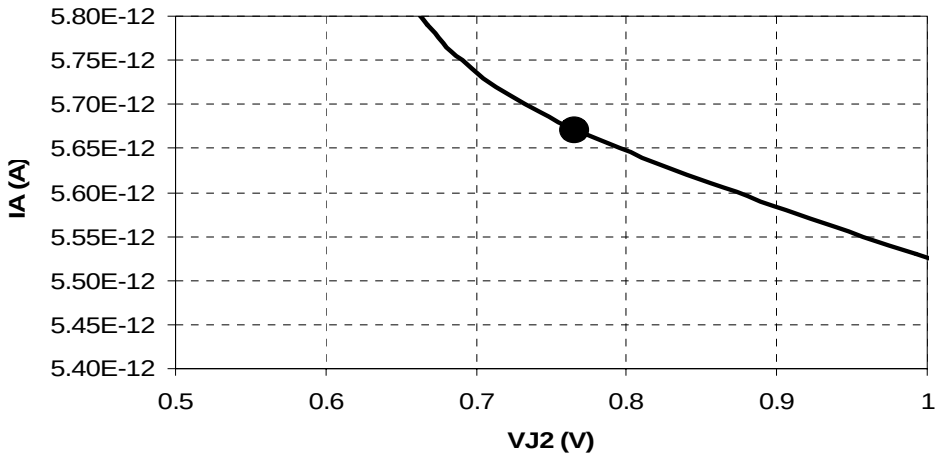

Figure 29. Anode Current vs. J2 Junction Potential.

Figure 30 shows the location of the holding current as determined above compared to the value determined by the vertical tangent. The graphical solution occurs at a much lower anode current, and therefore does not occur at the minimum applied voltage across the entire thyristor.

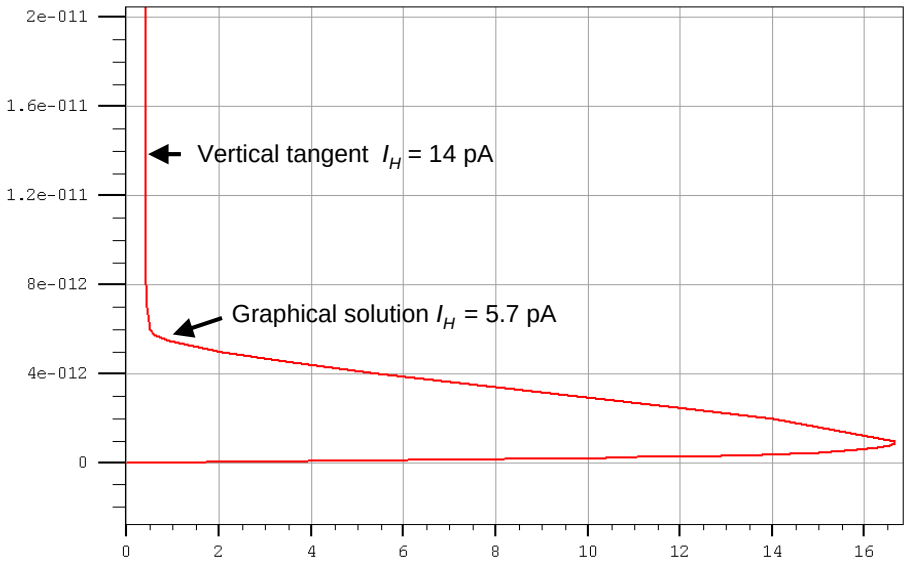

Figure 30. Close Up View of IV Curve.

4. Variation of parameters

This section discusses a series of simulations performed on the basic device structure already introduced, but with one of the key parameters varied while the others are

held constant. This will allow observation of the effect that parameter changes have on the IV curve. The device parameters that will be varied are the doping concentration of the two BJT emitters, P1 and N2, and thickness of the two BJT bases, N1 and P2. In each plot the IV curve for the device as already described will be given as a reference. It will be indicated in the legend of each figure by the word *reference* attached to the file name.

a. Variation of Emitter Doping

The P1 layer is described in the two-transistor model as the emitter for the *pnp* transistor and the N2 layer is the emitter for the *npn* transistor. Equation (2.3) defined the emitter efficiency γ of a BJT with the assumption that variations in γ were determined primarily by changes in the relative doping of the base and emitter. This section continues that assumption to explore the changes to the simulated IV curve as the doping levels are varied in the two BJT emitters. The doping levels were changed separately to allow observation of the effects on the switching voltages and holding currents independently.

Figure 31 shows the change in the IV curve as P1 layer doping is increased. At the low end of this range, emitter efficiency at J1 is very low, and approaches unity at the high end of the range. The legend in Figure 31 shows the name of the ATLAS output file which corresponds to the various doping concentrations in the P1 layer, which are given in scientific notation form. Thus, 1e15, shown in red represents the IV curve simulation result when P1 is doped to $1\times10^{15}\,\mathrm{cm}^{-3}$. This notation is used through this section when doping levels are varied.

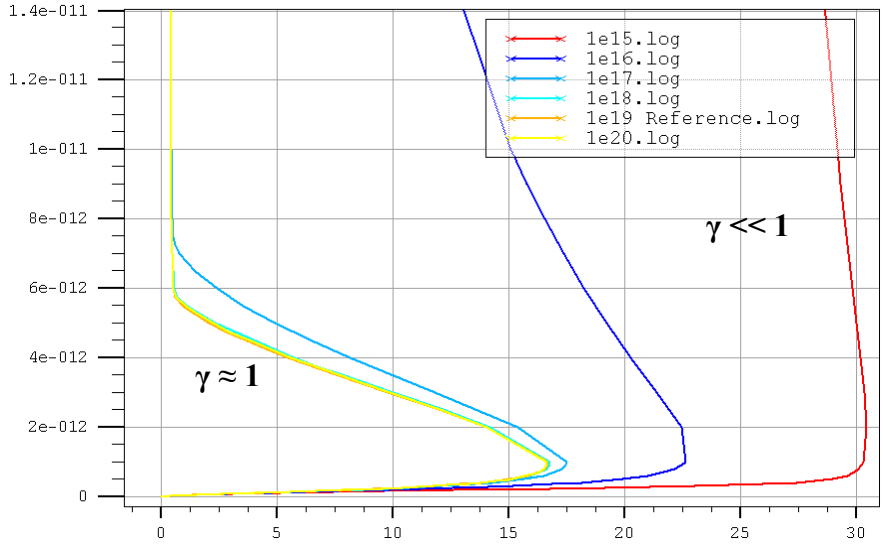

Figure 31. Anode Current vs. Bias as the P1 Doping Parameter is Varied.

Figure 31 shows that high emitter efficiency at J1 enhances the completion of the turning-on process by lowering the holding current. The lowest holding current occurs at the highest emitter efficiency. Decreasing the doping level decreases the emitter efficiency, which also increases the holding current. The lower doping also increases the switching voltage.

Figure 32 shows the same doping changes applied to the N2 layer. As with the previous case, lower N2 doping results in lower emitter efficiency at J3, higher holding current, and higher switching voltages. However, unlike the previous case, the lowest holding current does not occur at the highest emitter efficiency. This suggests that, in the case of N2 doping, there exists an optimal doping level to enhance thyristor turn-on. This optimum doping is not equal to the maximum feasible doping as in the previous case.

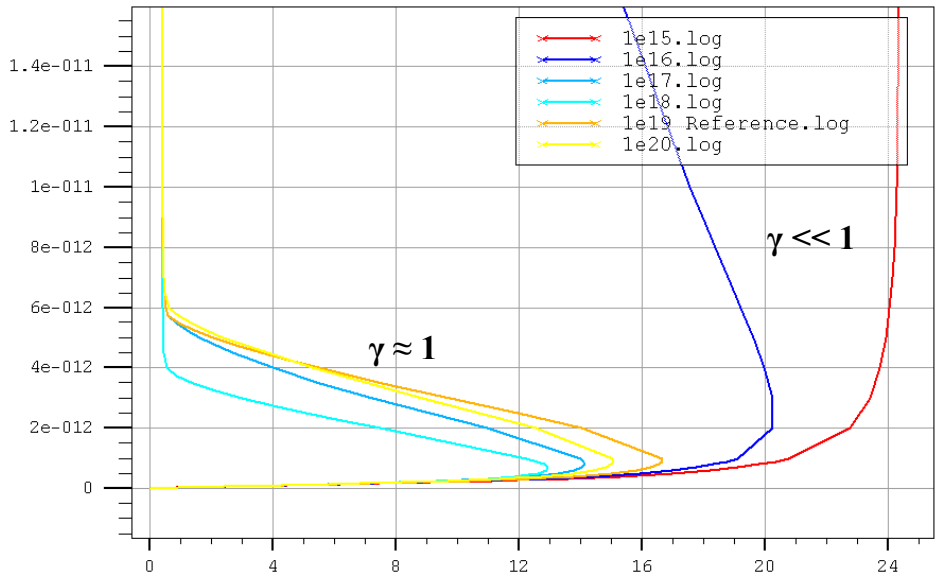

Figure 32. Anode Current vs. Bias as the N2 Doping Parameter is Varied.

As shown in Figures 31 and 32, the lowest emitter doping cases, shown in red, resemble the reverse-bias breakdown curve of a diode, which is caused by impact ionization. The higher bias at the switching point for the low emitter doping cases also means that impact ionization happens at a higher rate. Figure 33 shows an additional simulation performed using the N2 emitter doping level of $1\times10^{15}\,cm^{-3}$. In this figure, the red line represents the simulation with the impact model enabled and the blue line represents the simulation with the impact model disabled. This indicates, again, that impact ionization is not the dominant factor in thyristor operation.

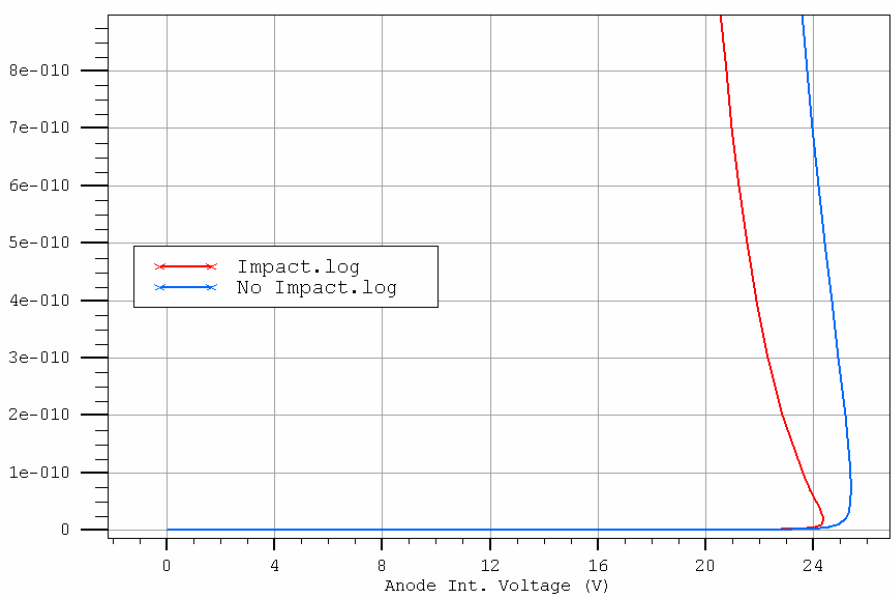

Figure 33. Two IV Curves With and Without the Impact Ionization Model Enabled.

The increased switching voltage for low emitter doping can be explained using the theory that switching happens when the sum of the current gains α_{npn}, α_{pnp} are equal to or greater than one. Low emitter doping will reduce the current gain, thereby increasing the switching current. In the lowest doping case for N2, the switching current is an order of magnitude higher than the higher doped cases. This indicates that low emitter efficiency is the primary explanation for the higher switching voltage and current for low emitter dopings.

From the consideration of designing a thyristor detector, Figures 31 and 32 suggest that lower emitter efficiency is desired to enhance the operation of a thyristor-type detector. This will increase the holding point to make it less likely that the turn-on process is completed. However, the increase in switching voltage may make the holding current improvement unacceptable.

b. Variation of Base Thickness and Doping

In this section the thicknesses and doping levels of the N1 and P2 layers, which are the bases of the *pnp* and *npn* transistors in the two-transistor model, respectively, will be varied. The IV curve as the P2 base thickness is varied is shown in Figure 33. Increases in P2 thickness have the predictable effect of increasing the switching

voltage and holding current. The basic shape of the IV curve, however, remains the same as P2 thickness varies. The legend for Figure 34 shows the names of the ATLAS output files which correspond with the simulation results for various layer thicknesses, which are specified in microns. Thus, 1_5.log corresponds to the simulated IV curve when the thickness of the P2 layer is 1.5 μm. This convention is used throughout this section the for layer thickness variation plots.

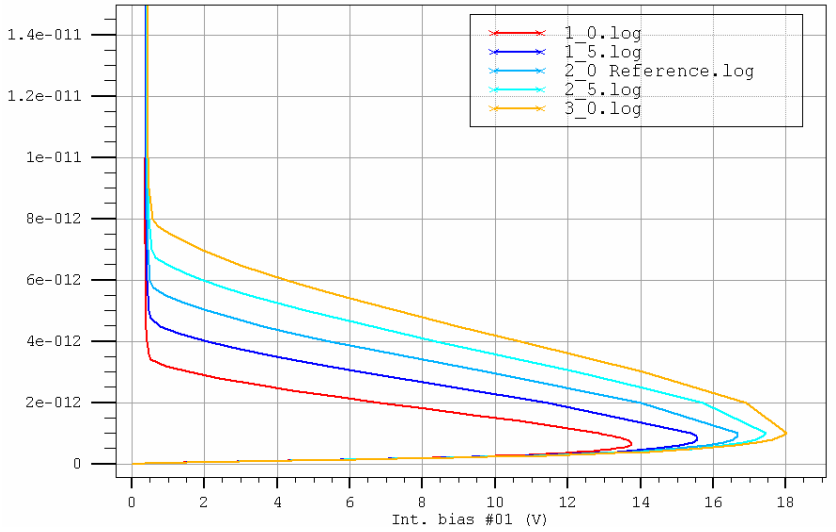

Figure 34. Anode Current vs. Bias as the P2 Thickness Parameter is Varied.

Figure 35 shows the changes to the IV curve as the P2 doping level is varied. The effect of increasing the doping level of P2 is similar to decreasing the doping level of N2; it reduces the emitter efficiency. Figure 35 shows, as does Figure 33, that lower emitter efficiency results in higher switching voltage and holding current. The increase in switching voltage with increased thickness can be explained by considering that fewer electrons injected through J3 will successfully transit the P2 when it is thicker. In BJT terms, a thicker P2 layer results in a lower base transport factor [4].

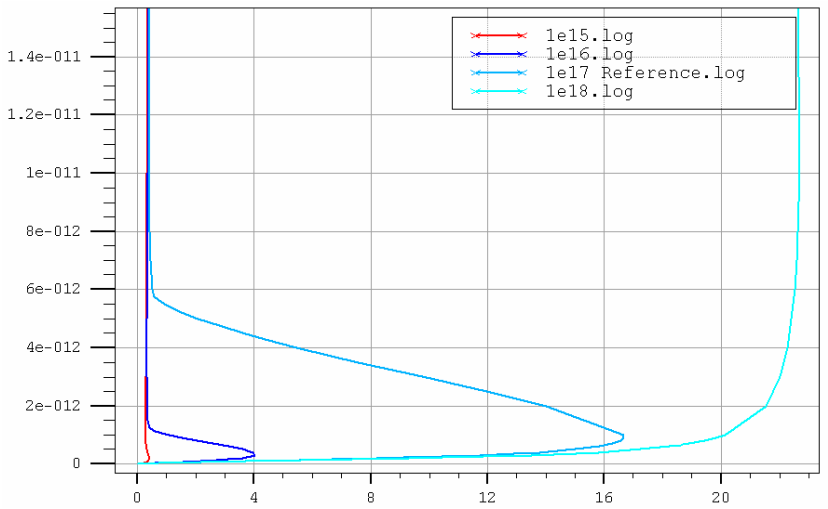

Figure 35. Anode Current vs. Bias as the P2 Doping Parameter is Varied.

Next, the same variations were performed on the N1 doping and thickness. Figure 36 shows the simulated IV curve as N1 thickness is varied. It is immediately evident that the operation of the thyristor is much more sensitive to changes in N1 than in P2. Notice that for IV curve associated with N1 thickness of 0.5 μm the device does not behave as a thyristor, but as a diode.

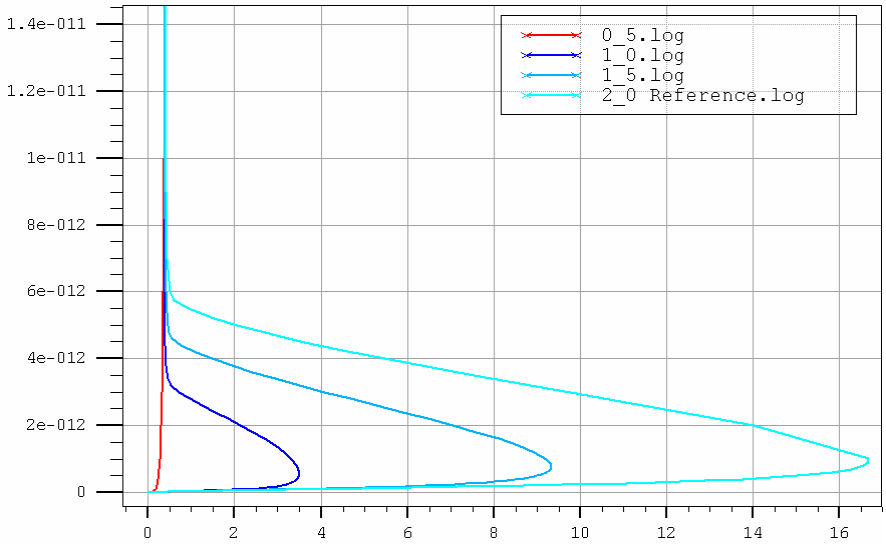

Figure 36. Anode Current vs. Bias as the N1 Thickness Parameter is Varied.

Figure 37 shows the change in the simulated IV curve as N1 doping is varied. Again, the switching voltage is strongly influenced by the changes in the doping level of N1.

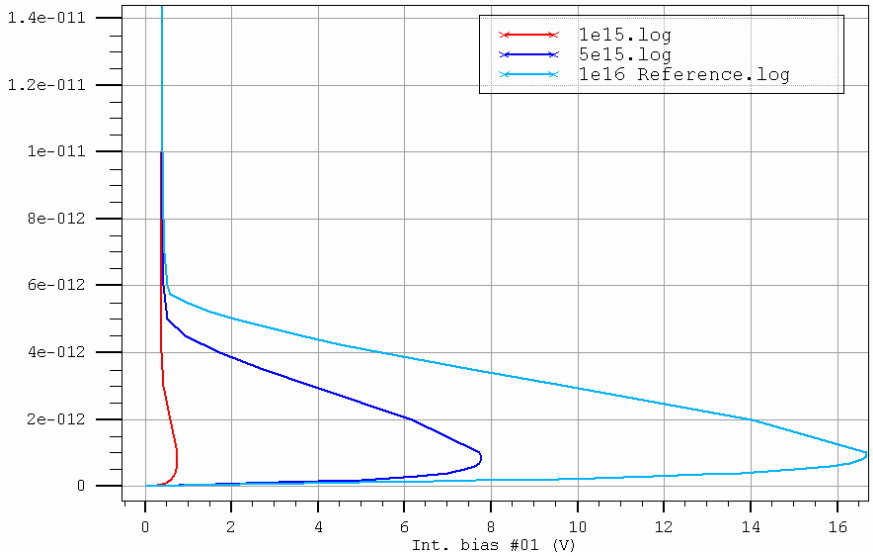

Figure 37. Anode Current vs. Bias as the N1 Doping Parameter is Varied

Figures 36 and 37 show that the thyristor switching point is primarily dependent on the doping and thickness of the N1 layer, with a secondary dependence on the P2 doping. These parameters can be used to determine the switching point for a thyristor, and the remaining parameters, including the doping of P1, P2 and N2, can be used to shape the IV curve as desired.

5. Optimal Thyristor Detector Parameters

The results of the previous section can be used to develop a few conclusions related to optimizing detector operation of a thyristor. These results are necessarily general relationships between layer parameters.

The first goal of a thyristor designed to produce a pulse-mode output is that it must not complete the turning-on process. That is, the anode current must not exceed the holding current. This can be achieved most effectively by lowering the efficiency of the emitters in the two-transistor model, P1 and N2. This can be done by decreasing the doping of the emitters, P1 and N2, or increasing the doping concentration of the P2 base.

Lesser effects can be achieved by increasing the thickness of the N1 or P2 layers, or increasing the doping concentration of the N1 layer.

The switching voltage can also be controlled by varying the parameters. The primary determinant in the switching voltage is the doping and thickness of the N1 layer. These parameters can be used to select a switching voltage slightly above the intended biasing voltage for the thyristor.

E. CHAPTER CONCLUSION

This chapter has introduced the thyristor device and described the static theory of thyristor operation. ATLAS simulations were introduced and applied to demonstrate thyristor operation in different operating modes. It was shown that the switching point voltage is determined primarily by the thickness and doping of the N1 layer and the doping of the P2 layer. The final section of the chapter showed how device parameters can be used to modify the shape of the IV curve. The following chapter will apply some of these items in the effort to design a thyristor to be used as an optical detector.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DEVICE DESIGN AND SIMULATION

This chapter discusses the design and simulation of the photodetecting thyristor. As originally proposed, this thesis would have stopped at the computer simulation stage. However, the author was awarded a research Fellowship from the Space and Naval Warfare Systems Center, San Diego, to extend the research effort to include fabrication of a device to test the concept. Because the late funding of this project, the design has been submitted for fabrication but the fabrication is not yet complete. This chapter presents an overview of the process used to create the device design, an introduction to the design created, and the results of simulations performed on the device. The device designed is not an ideal detector structure; rather it is a functional design which can be built using an existing fabrication process.

There are two procedures used to create the device design. The first is creating the device in ATLAS to optimize placement of the thyristor layers. This is done using the constraints of the fabrication process parameters and the layout design rules. The second procedure converts the design to a layout file which is used to specify how the device is to be fabricated. The second step was done using L-Edit by Tanner, Inc. Before the design could be started, it was necessary to select a fabrication process. The steps taken to select the process and design the device are described below.

A. PROCESS SELECTION AND PARAMETER ESTIMATION

There are very limited options available to fabricate small runs of integrated circuit devices at an affordable cost. Perhaps the most widely used fabrication service for small lots is provided by the MOSIS service. MOSIS is not a manufacturer; rather it is an organization that compiles multiple small-run projects submitted by companies, governments, and universities onto a single wafer which is then fabricated by a semiconductor fabrication company. MOSIS works with several different manufacturers allowing for a wide variety of available processes. The result is that MOSIS provides a method for getting small-lot projects manufactured using a specific process at a reasonable cost.

The considerations used in choosing a fabrication process include minimum feature size, cost, and the number and types of layers available to the circuit designer. This project does not involve particularly small features or excessive numbers of layers, but it does require that the process permit bipolar devices. The bipolar requirement proved to be the real limit in selecting the process. MOSIS offers only one such process, the 1.50-μm ABN CMOS process (ABN) by AMI Semiconductor, Inc. As with any commercial process, the circuit designer has no control over the doping concentration and thickness of the various layers. This required estimation of the process parameters to permit simulations.

1. The AMI ABN Process

The ABN process allows up to two metal layers, two polysilicon layers, an *n* well and a *p* base, all on a *p* substrate. This process was designed to permit a vertical *npn* BJT as shown in Figure 38.

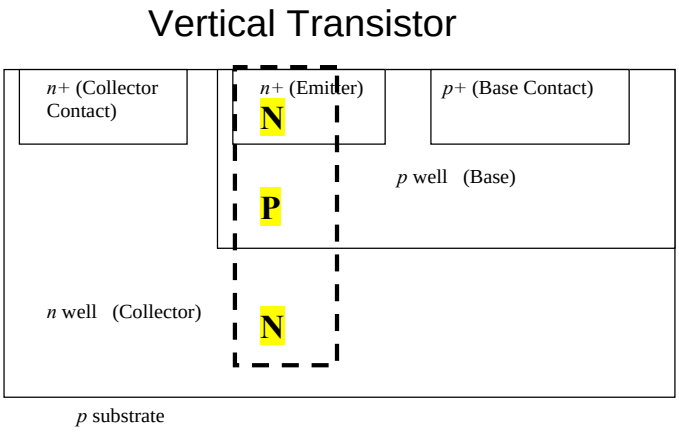

Figure 38. Cross-Section View of the ABN Process Vertical BJT Structure (After [25].).

This basic scheme was modified slightly to create the *pnpn* thyristor needed for this project by replacing the *p*+ collector contact with an *n*+ layer. This allowed the layers of the thyristor to be arranged as shown in Figure 39. Rather than a vertical BJT, the layout becomes a lateral thyristor.

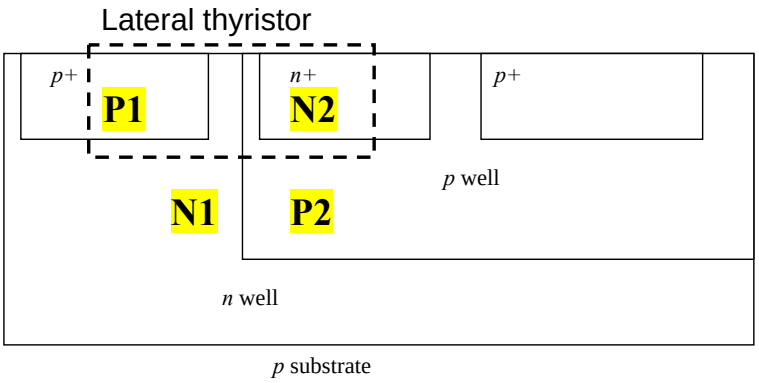

Figure 39. Cross-Section View of the Thyristor Structure using the ABN Process.

2. Parameter Estimation

One of the difficulties encountered using a commercial fabrication process is that the doping levels, doping profiles, and layer thicknesses are not available to the designer because they are considered proprietary data. The designer is then tasked to design a device based on fixed, unknown parameters. This is not usually a problem for most circuits since they typically use manufacturer-provided circuit elements. These are sub-circuits and devices which are known to be designed correctly to perform a specified operation. Then the circuit designer's job becomes connecting these known-good elements into a larger circuit to accomplish a task.

This project, however, is not based on these predetermined device layouts. Rather, the goal here is to create an entirely new device whose operational characteristics depend on the unknown parameters. Specifically, this includes the doping concentrations of all the layers and the thicknesses of the N1 and P2 regions. The importance of these parameters meant that estimates of the critical parameters were needed.

Doping concentration in the various layers can be derived by first determining the resistivity ρ of each region. This can be done by using

$$\rho = R_S x \,, \tag{3.1}$$

where R_S represents the sheet resistance of the material and x represents the thickness of the material. Sheet resistance is non-proprietary data which can be obtained from MOSIS

for previous manufacturing runs [25]. These values are expected to remain nearly constant for future fabrication runs. The value of R_S for the thyristor layers, averaged over the last twenty fabrication runs, is given in Table 1. To complete the calculation of resistivity ρ, the layer thickness x must be known or estimated. Since these thicknesses are not known, it was necessary to make estimates based on comparison with similar CMOS processes [26, 27]. Once the resistivities of the layers were estimated, the doping concentrations were estimated graphically using Figure 40. Estimated layer thickness, resistivity, and doping concentration are given for each of the layers used in the thyristor in Table 1.

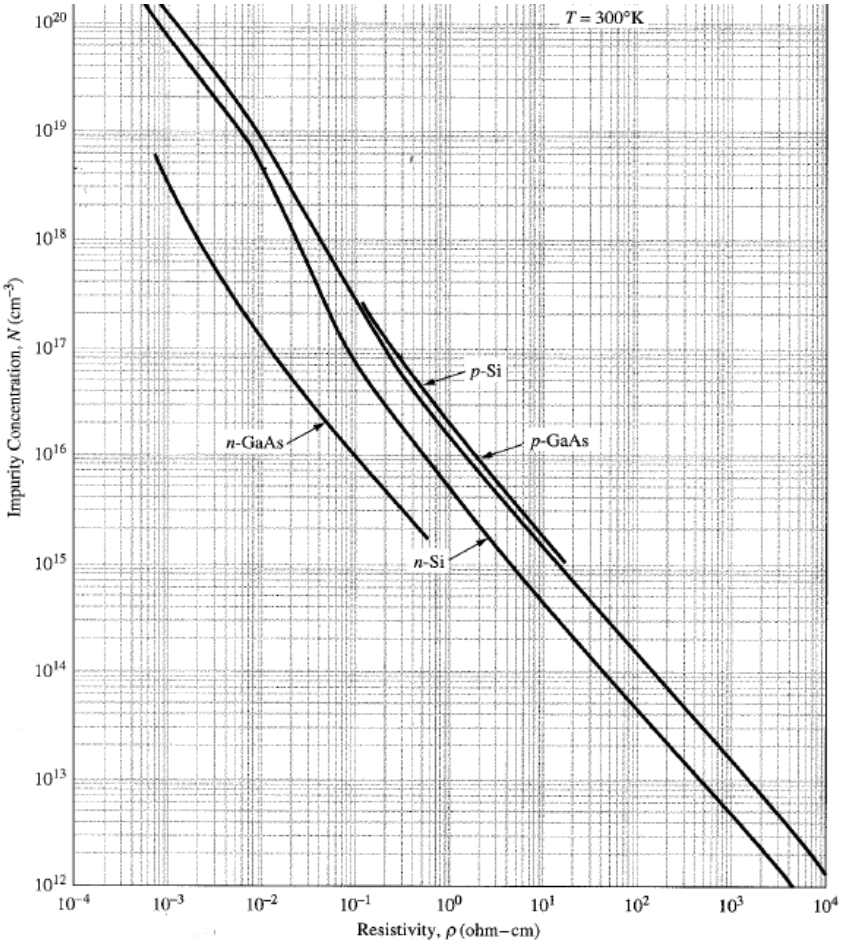

Figure 40. Resistivity vs. Impurity Concentration (From [14].).

Region	Average R_S (Ω/square)	Estimated Thickness (μm)	Resistivity $(\Omega\cdot\text{cm})$	Doping (cm^{-3})
P1	76.5	0.3	0.0022	4×10^{19}
N1	1598	4	0.64	7×10^{15}
P2	2265	2	0.45	2×10^{16}
N2	52.5	0.3	0.0016	4×10^{19}

Table 1. Critical Parameters of Device Layers.

B. DEVICE DESIGN

Before the ATLAS model could be created it was necessary to understand the manufacturer's restrictions and limitations on dimensions and placement of the various layers for the process chosen. These restrictions are specified as *design rules*. This section describes how the basic device was created within the restrictions imposed by the design rule set.

1. Design Rule Non-Disclosure Agreement

The actual design rules for the ABN process are protected by a non-disclosure agreement with the manufacturer. Therefore the actual dimensions of the device as submitted for fabrication will not be given in this thesis. The simulations and results described in this chapter were obtained using parameters that are not derived from the protected data. Rather, the design is presented using the design rules for the scalable CMOS (SCMOS) process, which are available on the MOSIS website [25].

2. SCMOS Design Rules

This section introduces the design rules used to create the simulation of the thyristor device. As discussed above, these are not the design rules used for the actual fabrication of the device because of the non-disclosure agreement, rather they are the SCMOS rules. Only a small portion of the entire SCMOS rule set will be introduced. Specifically, only those rules needed to describe the device for simulation using ATLAS are discussed. This excludes the rules for contacts, metal layers, interconnects, and bonding pads since these are not used for the simulation. The full set of SCMOS design rules is available on the MOSIS web site [25].

The primary feature of this design is a thyristor structure created using the basic structure of the *p* base layer inside the *n* well layer. Figure 41 shows a graphical repre-

sentation of the design rules pertaining to these layers. The numbers in Figure 40 are references to the rules which specify minimum dimensions allowable in the SCMOS rule set, which are listed in Table 2. The units in Table 2 are given in terms of the unit *lambda*. Lambda is a parameter used in SCMOS to allow uniform scaling of all device dimensions by the same factor. This allows the same design layout to be used in different technologies without resizing the layout. The layout for this project was not drawn using lambda-based rules but some value is needed to convert the SCMOS rules to a device which can be simulated. For the purposes of this chapter, lambda will be equal to 0.8 μm. The design rule set will be used as shown, except for the N+ Select region near the top. As explained earlier, this will be replaced by a P+ Select layer to obtain the *pnpn* thyrsitor structure.

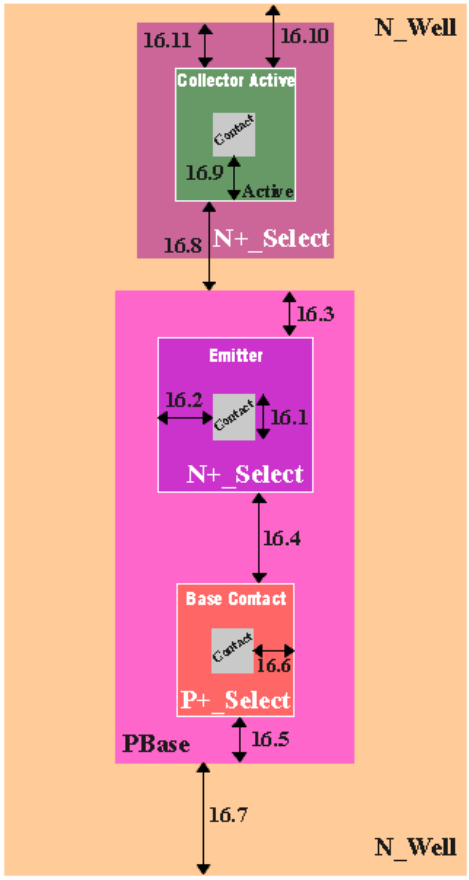

Figure 41. SCMOS Design Rules Diagram for *N* Well and *P* Base Layers (Top View) (From [25].).

Rule	Description	Lambda
16.1	Contact	2 x 2
16.2	Minimum N+_Select overlap of contact	3
16.3	Minimum PBase overlap of N+_Select	2
16.4	Minimum spacing between N+_Select and P+_Select	4
16.5	Minimum PBase overlap of P+_Select	2
16.6	Minimum P+_Select overlap of Contact	2
16.7	Minimum N_Well overlap of PBase	6
16.8	Minimum spacing between PBase and Active	4
16.9	Minimum Active overlap of Contact	2
16.10	Minimum N_Well overlap of Active	3
16.11	Minimum P+_Select overlap of Active	2

Table 2. SCMOS Design Rules for *N* Well and *P* Base Layers (After [25].).

3. ATLAS Input Deck Creation

Once the design rules are known, they were then used to create the cross-section of the device to be used in ATLAS simulations. This section discusses the layers used in the ABN process, which are shown in Figure 41, and how they are used to create the thyristor device shown in Figure 42.

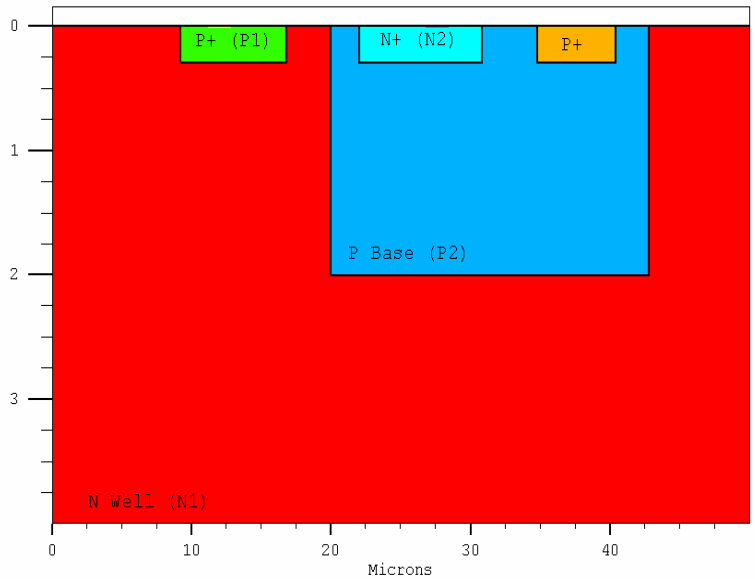

Figure 42. Thyristor Cross-Section View.

a. Substrate

The substrate used in this device is doped lightly *p* type silicon. It is not used in the ATLAS simulations since it is assumed to provide isolation between adjacent devices on the actual chip. The substrate is not shown in either Figure 41 or 42.

b. N Well

The N_Well contains the entire thyristor device structure. It also acts as the N1 layer in the thyristor device. As discussed earlier, the doping level is estimated to be 7×10^{15} cm^{-3}, and the layer thickness is estimated to be 4.0 μm. The width of the *n* well, shown in red in Figure 42, is 50 μm, while the length is either 110 μm or 510 μm.

c. PBase

The P Base layer, shown in dark blue in Figure 42, serves as the P2 layer for the thyristor. The doping level is estimated to be 2×10^{16} cm^{-3} and the layer thickness is estimated to be 2.0 μm. The P Base layer serves as the P2 layer for the thyristor. It also contains the contact to the P2 layer, shown in Figure 41 as P+ Select, and the N2 layer, shown as N+ Select. The width of the P Base layer 22.8 μm. Design rule 16.7 specifies that the P Base must be at least 6 lambda $(4.8\ \mu\text{m})$ from the edge of the N_Well edge.

d. N+ Select

The N+ Select layer, shown in light blue in Figure 42, is used only for the N2 layer in the thyristor. The doping level of this layer is estimated to be 4×10^{19} cm^{-3}, and the layer thickness is estimated to be 0.3 μm. The cathode is attached to this layer by a Contact layer. Design rule 16.1 states that the contact must be 2 lambda square and rule 16.2 states that the contact must be 3 lambda from the edge of the N+ Select layer. The same contact size is used in all regions in this process.

Because light is detected by being absorbed in the junction formed by the N Well and the P Base layers, it is important that the contacts and the overlying metal layers be placed as far away from the junction as possible. This was accomplished by making the N+ Select layer wider than the minimum width and moving the contact away from the junction. This technique is also applied to the cathode contact to the P1 layer.

e. P+ Select and Active

The P+ Select layer is used to provide the gate contact to the P2 thyristor layer, which is formed by the P Base layer, and to create the P1 layer. The doping level of this layer is estimated to be 4×10^{19} cm^{-3} and the layer thickness is estimated to be 0.3 μm.

The P+ Select region used for the gate contact is located inside the P Base layer. It is shown in orange in Figure 42. Design rule 16.4 states that the P Base must be placed at least 4 lambda from the N+ Select layer and rule 16.6 states that it must be placed at least 2 lambda from the edge of the P Base. The standard contact size is used and it must be placed at least 2 lambda from the edge of the P+ Select.

The P+ Select region used to create the P1 region requires a more careful description. The P+ region, shown in green in Figure 42, actually corresponds with the Active layer in the ABN process. As shown in Figure 41, the Active layer lies within the P+ Select layer. The ion implantation process, which is the method of doping the region, only takes place in the active area. In other words, the portion of the P+ Select region outside of the Active region retains its original doping. In this case, it retains the doping of the N Well region and only the Active region is doped at the P+ level of 4×10^{19} cm^{-3}. This can be seen in Figure 42 as the gap between the P+ Select and the P Base. This gap then defines the width of the N1 layer, which is the primary factor in determining the switching point of the thyristor. In the case of the SCMOS design rule, the minimum N1 width is 4 lambda, or 3.2 μm. This represents the basic device design.

Because of the relatively large chip area available and the small size of the device, three other variations were also designed, both for comparative simulations and for fabrication. This allows multiple devices to be built on a single chip. The N1 widths of the variant designs are 4.2 μm, 5.2 μm, and 2.2 μm. The last N1 width listed violates the design rules for the process. This may mean that the fabricated device may not perform as expected, but this will not affect the operation of the other devices on the chip.

The design rules which apply to the Active region include the following; 16.8, which specifies a minimum separation from the P Base of 4 lambda; 16.9, which

specifies the minimum spacing of the contact from the Active edge of 2 lambda; 16.10, provides minimum spacing of the Active from the edge of the N Well of 3 lambda; and 16.11, which provides minimum spacing of the Active from the edge of the P+ Select of 2 lambda.

4. ATLAS Simulation Results

The design rules discussed in the previous section were used to create an ATLAS input deck to simulate the IV curve. The structure for this device was shown in Figure 42. Figure 43 shows the IV curve for the four variations of the basic device, with the N1 width indicated in microns in the legend. As expected, the structures with the wider N1 width have a higher switching voltage.

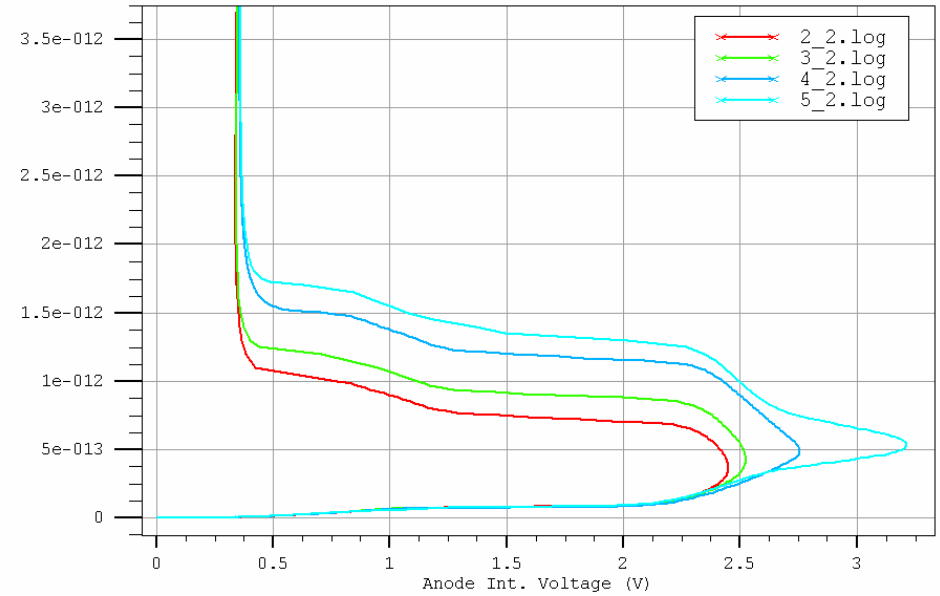

Figure 43. Simulated IV Curve of Fabricated Device.

Figure 44 shows a structure plot displaying the simulated electric field in the 3.2 μm device biased at the switching threshold. The peak electric field magnitude at J2 is 65,000 V/cm. According to Figures 24 and 25, impact ionization occurs at a very low rate.

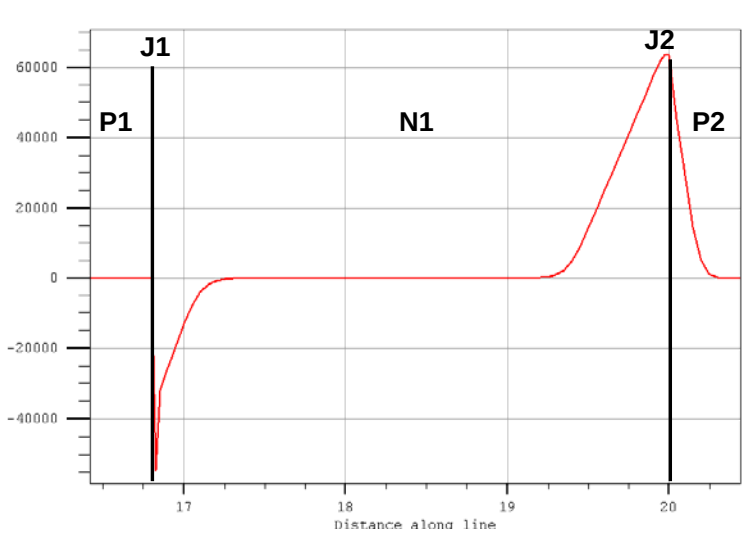

Figure 44. Electric Field in N1 When Biased at the Switching Threshold.

C. DEVICE LAYOUT

Once the design was verified using ATLAS simulations, it had to be converted to a useable layout for fabrication. This was done using L-Edit by Tanner, Inc. Figure 45 shows the top view of the basic device layout. The cross section view shown in Figure 42 could be taken along a vertical line through Figure 45. The ABN process layers are pointed out in this figure, along with the thyristor layer names. The length of the basic device, which would be along a horizontal line in Figure 45, was chosen to be 110 μm.

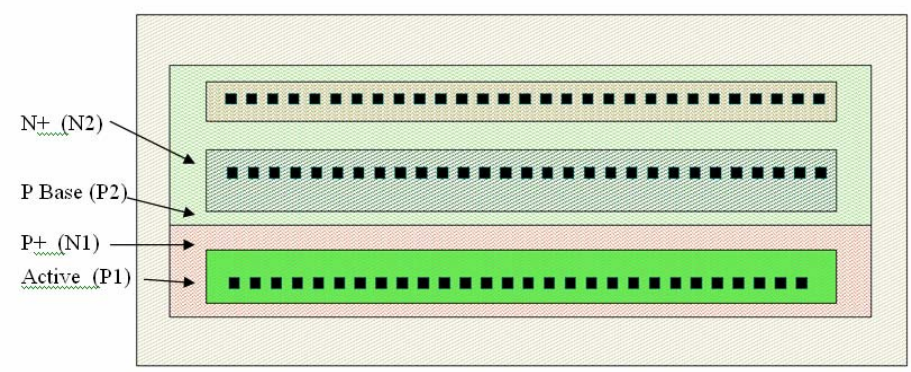

Figure 45. Top View of Device Layout.

The surface area of the chip produced by the ABN process is 2.2 mm square. This, of course, is much larger than the area of the device in Figure 45. This allowed

multiple devices with different lengths to be placed on the chip. The chip design submitted for fabrication consists of sixteen separate devices. Each of the four N1 width variants are designed with two lengths, 110 μm and 510 μm, for a total of eight distinct device designs. Two copies of each of these eight designs are included on the chip. A top-view of the final chip design is shown in Figure 46.

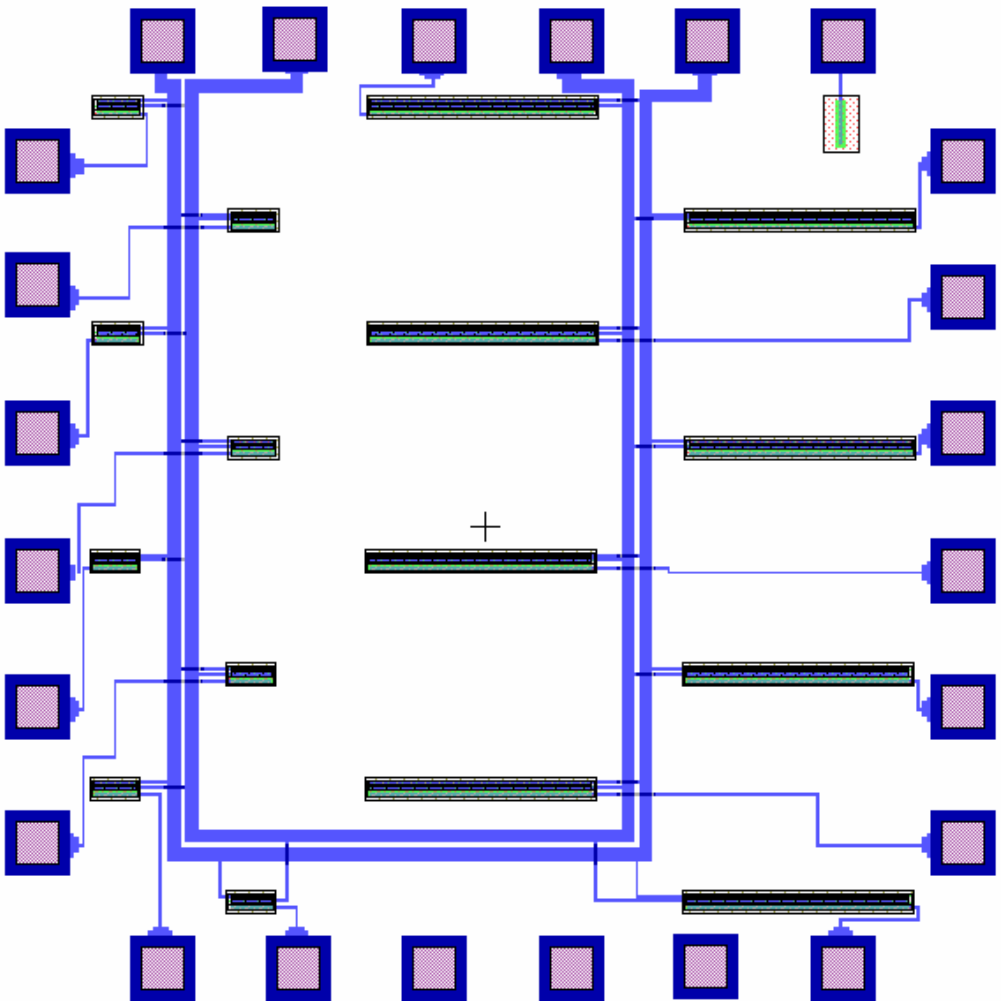

Figure 46. Layout of Chip as Submitted for Fabrication.

D. CHAPTER CONCLUSION

This chapter introduced the design of the device created and the methods used to create that design. ATLAS simulations were used to validate the operation of the device

using the principles introduced in Chapter II. The following chapter briefly summarizes the work performed for this thesis and makes recommendations for further research.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. CONCLUSIONS AND RECOMMENDATIONS

This thesis has conducted an in-depth investigation into static thyristor operation. The emphasis of this work was to develop a better understanding of thyristor operation to enable further work toward designing a thyristor structure for use as a pulse-mode optical detector. A key result is the identification of desirable doping concentration relationships between the layers needed to lower the emitter efficiency, thereby making the thyristor more resistant to completely turning on.

This thesis is the first of several research efforts needed to advance the idea of a thyristor-based detector. The next logical step will be to test and analyze the thyristor device described in Chapter III. The previous experiments using a thyristor as a detector [2], described in Chapter I, were performed on a device with many unknowns. This includes not only the layer structure and doping concentrations of the thyristor but also how the device actually sits in the package. Those experiments were necessarily limited to qualitative results. Although the new device is not an ideal detector structure, and some of the parameters can only be estimated, meaningful quantitative results which can be used to improve the structure are expected.

Another area of research is extending the analysis of thyristor operation into the dynamic regime. Of particular interest are the physical changes in the device which produce the pulsed output. It is expected that the change in junction capacitance as current increases is strongly involved in the pulse generation. ATLAS has the capability to perform transient simulations, so should continue to be of use in this effort.

Finally, work needs to be done to develop an optimal detector structure. The device designed in this thesis has many shortcomings as a detector because of the extremely limiting design rules for the process used. Other fabrication processes may have less restrictive rules which will permit a larger area for useful optical absorption.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. ATLAS INPUT DECKS

The following input deck code was used to generate the simulated IV curve for the basic device as described in Chapter III.

```
GO ATLAS

# Sets variable used to input the doping concentrations for the thyristor lay-
ers
SET P1_DOPING=4.0e19
SET N1_DOPING=7.0e15
SET P2_DOPING=2.0e16
SET N2_DOPING=4.0e19


# Defines variables to be used to define to corners of the various layers
SET X01=07.2
SET X02=09.2
SET X03=11.2
SET X04=12.8
SET X05=16.8
SET X06=20.0
SET X07=20.0
SET X08=22.0
SET X09=26.8
SET X10=28.4
SET X11=30.8
SET X12=34.8
SET X13=37.2
SET X14=38.8
SET X15=40.4
SET X16=42.8
SET X17=50.0
SET Y01=0.3
SET Y02=2
SET Y03=4

# Defines the mesh used carry out the numerical solution by ATLAS.
MESH
X.MESH LOC=0.0 SPACING=1
X.MESH LOC=$X01 SPACING=0.5
X.MESH LOC=$X02 SPACING=0.5
X.MESH LOC=$X03 SPACING=0.2
X.MESH LOC=$X04 SPACING=0.1
X.MESH LOC=$X05 SPACING=0.05
X.MESH LOC=$X07 SPACING=0.05
X.MESH LOC=$X08 SPACING=0.05
X.MESH LOC=$X09 SPACING=0.1
X.MESH LOC=$X10 SPACING=0.25
X.MESH LOC=$X11 SPACING=0.5
X.MESH LOC=$X12 SPACING=0.5
X.MESH LOC=$X13 SPACING=0.5
X.MESH LOC=$X14 SPACING=0.5
X.MESH LOC=$X15 SPACING=0.5
X.MESH LOC=$X16 SPACING=0.75
X.MESH LOC=$X17 SPACING=1
```

```
Y.MESH LOC=0.0 SPACING=0.05
Y.MESH LOC=$Y01 SPACING=0.05
Y.MESH LOC=1 SPACING=0.1
Y.MESH LOC=2.25 SPACING=0.25
Y.MESH LOC=$Y03 SPACING=0.5

# Reduces mesh density in non-critical areas to speed up simulation
ELIMINATE Y.DIR Y.MIN=2.25 Y.MAX=$Y03 X.MIN=0 X.MAX=$X17
ELIMINATE Y.DIR Y.MIN=0.5 Y.MAX=2 X.MIN=0 X.MAX=14
ELIMINATE Y.DIR Y.MIN=2.5 Y.MAX=3.75 X.MIN=10 X.MAX=25

# Defines the regions in ATLAS, which are equivalent to the layers in a thyris-
tor.
#      Region 1 is the N1 layer
#      Region 2 is the P1 layer
#      Region 3 is the P2 layer
#      Region 4 is the N2 layer
REGION NUMBER=1 MATERIAL=SILICON
REGION NUMBER=2 MATERIAL=SILICON Y.MIN=0 Y.MAX=$Y01 X.MIN=$X02 X.MAX=$X05
REGION NUMBER=3 MATERIAL=SILICON Y.MIN=0 Y.MAX=$Y02 X.MIN=$X07 X.MAX=$X16
REGION NUMBER=4 MATERIAL=SILICON Y.MIN=0 Y.MAX=$Y01 X.MIN=$X08 X.MAX=$X11
REGION NUMBER=5 MATERIAL=SILICON Y.MIN=0 Y.MAX=$Y01 X.MIN=$X12 X.MAX=$X15

# Defines the location of the contacts to certain layers
ELECTRODE NAME=anode TOP X.MIN=$X03 X.MAX=$X04
ELECTRODE NAME=cathode X.MIN=$X09 X.MAX=$X10
ELECTRODE NAME=gate X.MIN=$X13 X.MAX=$X14

# Defines the doping profile of the layers
DOPING UNIFORM N.TYPE REGION=1 CONCENTRATION=$N1_DOPING
DOPING UNIFORM P.TYPE REGION=2 CONCENTRATION=$P1_DOPING
DOPING UNIFORM P.TYPE REGION=3 CONCENTRATION=$P2_DOPING
DOPING UNIFORM N.TYPE REGION=4 CONCENTRATION=$N2_DOPING
DOPING UNIFORM P.TYPE REGION=5 CONCENTRATION=$P1_DOPING

# Refines the grid is areas with rapoid changes in doping with distance
REGRID DOPING LOG RATIO=2 SMOOTH=4 X.MIN=$X04 X.MAX=$X09 Y.MIN=0 Y.MAX=$Y02
SOLVE INIT

# Saves and plots the structure file
save outf=A_3_2.str
tonyplot A_3_2.str

# Defines the mobility and recombination models used for this simulation
MODEL BIPOLAR FERMI

# Specifies that this simulation will be controlled by current through the an-
ode
CONTACT name=anode current

# Specifies the numerical technique used
METHOD NEWTON

# Opens the output file to store data generated by the following set of solu-
tion points
LOG OUTFILE=A_3_2.log

# Specifies the current levels where solutions are required
SOLVE INIT
solve ianode=1e-19
solve ianode=1.2e-19
solve ianode=1.4e-19
```

```
solve ianode=1.6e-19
solve ianode=2e-19
solve ianode=4e-19
solve ianode=1e-18
solve ianode=1e-17
solve ianode=2e-17
solve ianode=4e-17
solve ianode=7e-17
solve ianode=1e-16 istep=0.2e-16 ifinal=1e-15 name=anode
solve ianode=1.2e-15 istep=0.1e-15 ifinal=1e-14 name=anode
solve ianode=1.2e-14 istep=0.1e-14 ifinal=1e-13 name=anode
solve ianode=1.2e-13 istep=0.1e-13 ifinal=1e-12 name=anode
solve ianode=1.2e-12 istep=0.1e-12 ifinal=1e-11 name=anode
solve ianode=1.2e-11 istep=0.1e-11 ifinal=1e-10 name=anode

# Plots the IV curve
tonyplot A_3_2.log

quit
```

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

1. E. L. Dereniak and G. B. Boreman, *Infrared Detectors and Systems*, John Wiley & Sons, New York, 1996.

2. Gamani Karunasiri, Patent applied for in 2001.

3. D. Ziegler, P. Linderholm, M. Mazza, S. Ferazzutti, D. Bertrand, A. M. Ionecsu, and Ph. Renaud, "An active microphotodiode array of oscillating pixels for retinal stimulation," *Sensors and Actuators A: Physical*, Vol. 110, pp 11-17, 2004.

4. Robert F. Pierret, *Semiconductor Device Fundamentals*, Addison Wesley Longman, Reading, MA, 1996.

5. S. M. Sze, *Physics of Semiconductor Devices 2nd Edition*, John Wiley & Sons, New York, 1981.

6. V. A. K. Temple, "Comparison of light triggered and electrically triggered thyristor turn-on," *IEEE Trans. Electron Devices*, Vol. ED-28, No. 7, pp. 860-865, 1981.

7. Adolph Blicker, *Power Thyristor Physics*, Springer-Verlag, New York, 1976.

8. Sorab K. Ghandhi, *Semiconductor Power Devices*, John Wiley & Sons, New York, 1977.

9. William Shockley, *Electrons and Holes in Semiconductors With Applications to Transistor Electronics*, D. Van Nostrand, Toronto, 1950.

10. J. L. Moll, M. Tanenbaum, J. M. Goldey, and N. Holonyak, "*P-N-P-N* transistor switches," *Proc. IRE*, Vol. 44, pp. 1174-1182, 1956.

11. James F. Gibbons, "A critique of the theory of p-n-p-n devices," *IEE Trans. Electron Devices*, Vol. 11, pp. 406-413, 1964.

12. J. J. Ebers, "Four-terminal *p-n-p-n* transistors," *Proc. IRE*, Vol. 40, pp. 1361-1364, 1952.

13. E. S. Yang and N. C. Voulgaris, "On the variation of small-signal alphas of a p-n-p-n device with current," *Solid State Electronics*, Vol. 10, pp. 641-648, 1967.

14. H. Craig Casey Jr., *Devices for Integrated Circuits: Silicon and III-V Compound Semiconductors*, John Wiley & Sons, New York, 1999.

15. *ATLAS User's Manual*, vols. 1-2, Silvaco International, Santa Clara, CA, 2003.

16. Tibor Grasser, Ting-Wei Tang, Hans Kosina and Siegfried Selberherr, "A review of hydrodynamic and energy transport models for semiconductor device simulation," *Proc. IEEE,* Vol. 91, No. 2, pp. 251-274, 2003.

17. Mark Lundstrom, *Fundamentals of Carrier Transport 2nd Edition*, Cambridge University Press, Cambridge, 2000.

18. J. J. Liou, "Semiconductor device physics and modeling. 1. Overview of fundamental theories and equations," *IEE Proceedings G Circuits, Devices, and Systems*, Vol. 139, No. 6, pp. 646-654, 1992.

19. Mark Lundstrom, "Assumptions and trade-offs in device simulation programs," *Proc. 1992 Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 35-41, IEEE Press, New York, 1992.

20. Mark E. Law, E. Solley, M. Liang, and Dorothea E. Burk, "Self-consistent model of minority-carrier lifetime, diffusion length, and mobility," *IEEE Electron Device Let ters*, Vol. 12, No. 8, pp. 401-403, 1991.

21. Martin A. Green, "Intrinsic concentration, effective densities of states, and effective mass in silicon," *Journal of Applied Physics*, Vol. 67., No. 5, pp. 2944-2954, 1990.

22. J.W. Slotboom and H. C. De Graaf, "Measurements of bandgap narrowing in Si bipolar transistors," *Solid-State Electronics* Vol. 19, pp. 857-862, 1976.

23. Frank L. Pedrotti and Leno S. Pedrotti, *Introduction to Optics*, 2nd Ed., Prentice Hall, Upper Saddle River, NJ., 1993.

24. W. Maes, K. de Meyer, and R. Van Overstraeten, "Impact ionization in silicon: a review and update," *Solid-State Electronics*, Vol. 33, No. 6, pp. 705-718, 1990.

26. MOSIS ABN Process Webpage. http://www.mosis.org/Technical/Processes/proc-ami-abn.html Last accessed May 2005.

26. Tower Semiconductor Webpage http://www.towersemi.com/pages/products_content_pages.asp?intGlobalId=49&intLevel=2&intDeep=6 Last accessed May 2005.

27. Jong Duk Lee, Hyuck In Kwon, Jung Hyun Nam, Byung Chang Shim, and Byung Gook Park, "Design of One-Chip FED on a Standard CMOS Process," *Proc. 1st International Meeting in Information Display*, pp. 543-546, Taegu, Korea, August 2001.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Chair, Department of Electrical and Computer Engineering, Code EC
   Naval Postgraduate School
   Monterey, California

4. Professor Gamani Karunasiri, Code PH/Kg
   Naval Postgraduate School
   Monterey, California

5. Professor Douglas Fouts, Code EC/Fs
   Naval Postgraduate School
   Monterey, California

6. Marine Corps Representative
   Naval Postgraduate School
   Monterey, California

7. Director, Training and Decuation, MCCDC, Code C46
   Quantico, Virginia

8. Director, Marine Corps Research Center, MCCDC, Code C40RC
   Quantico, Virginia

9. Marine Corps Tactical Systems Support Activity (Attn: Operations Officer)
   Camp Pendleton, California

10. Major D. A. Moore
    Department of Electrical Engineering
    United States Naval Academy
    Annapolis, Maryland

THIS PAGE INTENTIONALLY LEFT BLANK